EUROPEAN JOURNAL OF CANCER 47 (2011) 1962–1970

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

CYP1A1*2A polymorphism as a predictor of clinical outcome in advanced lung cancer patients treated with EGFR-TKI and its combined effects with EGFR intron 1 (CA)n polymorphism

Qiang Nie, Xue-ning Yang, She-juan An, Xu-chao Zhang, Jin-ji Yang, Wen-zhao Zhong, Ri-qiang Liao, Zhi-hong Chen, Jian Su, Zhi Xie, Yi-long Wu *

Guangdong Lung Cancer Institute, Guangdong General Hospital and Guangdong Academy of Medical Sciences, Guangzhou, Guangdong, People's Republic of China

ARTICLE INFO

Article history:
Received 11 February 2011
Received in revised form 7 April 2011
Accepted 12 April 2011
Available online 26 May 2011

Keywords:
CYP1A1 polymorphisms
EGFR polymorphisms
EGFR mutations
EGFR target treatment
Combined effects

ABSTRACT

Background: Mutations in the epidermal growth factor receptor (EGFR) have been confirmed as predictors of efficacy for EGFR-tyrosine kinase inhibitors (TKIs). We investigated whether polymorphisms of the EGFR and cytochrome P450, family 1, member A1 (CYP1A1) genes were associated with clinical outcome in NSCLC patients treated with EGFR-TKI.

Methods: Genotypes for the intron 1 (CA)n repeat and R497K polymorphisms in the EGFR gene and the *2A (3801 T → C) and *2C (2455 A → G) polymorphisms in CYP1A1 gene were evaluated in 115 NSCLC patients by PCR-RFLP and DNA sequencing. Genetic polymorphisms were correlated with clinical outcomes of EGFR-TKIs. From a subgroup of patients whose tumour tissues were available, associations between somatic EGFR mutations, EGFR expression, and genomic polymorphisms were also analysed.

Results: EGFR intron 1 (CA)n and CYP1A1*2A polymorphisms were independent predictive factors ($p = 0.046$, $p = 0.011$, respectively) and the latter was also a prognostic factor ($p = 0.001$) for patients treated with EGFR-TKIs. We also observed a strong synergistic effect from two genotypes. Specifically, patients with both the T/T allele of the CYP1A1 gene and shorter intron 1 CA repeats ($\leqslant$16 CA) of the EGFR gene showed an improved response ($p = 0.002$) compared with patients with the T/C or C/C allele and longer intron 1 CA repeats (both alleles >16 CA). In contrast, for R497K and CYP1A1*2C, no relationship was observed with clinical outcome for patients treated with EGFR-TKIs ($p = 0.573$; $p = 0.629$, respectively). Both SNPs in the CYP1A1 gene showed a correlation with EGFR somatic mutations.

Conclusions: The findings of this study suggest that the CYP1A1*2A polymorphism is a predictor for clinical outcome in NSCLC patients treated with EGFR-TKI therapy, and combining analysis of both CYP1A1*2A and EGFR intron 1 (CA)n polymorphisms may be useful for predicting treatment outcome in NSCLC patients treated with EGFR-TKIs.

1. Introduction

EGFR-TKIs, such as gefitinib or erlotinib, which target EGFR-I,[1] have become standard second-line therapies for advanced non-small-cell lung cancer (NSCLC).[2] They are less toxic to normal cells and have improved tolerability in patients. Nevertheless, these targeted therapies have modest activity when given to unselected patient populations. However, there is

* *Corresponding author:* Fax: +86 20 83844620.
E-mail address: syylwu@live.cn (Y.-l. Wu).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2011.04.018

increasing evidence that the EGFR-TKIs have improved efficacy when prescribed to a 'targeted group' of patients. Female patients, non-smokers, and patients with a histological diagnosis of adenocarcinomas have been observed more frequently to respond to EGFR-TKIs.[3,4] These clinical characteristics are strongly associated with somatic mutations of the EGFR gene, which are known to be associated with a high tumour response.[5–7] Moreover, studies have confirmed that Asian patients with NSCLC are more responsive to EGFR-TKIs than Western populations and EGFR mutations are more common in this patient group.[8,9] In the Iressa Pan-Asia Study (IPASS), 59.7% of tumours in a clinically selected population had EGFR mutations, and the response rate to EGFR-TKIs was approximately 77–82%.[10,11] Thus, genetic factors seem to have notable relevance in the area of targeted anticancer therapy, and this has prompted efforts to determine the individual factors that predict such susceptibility. However, many patients have advanced and inoperable cancer at diagnosis, and it is not always possible to obtain tumour tissues to analyse for EGFR mutations; however, polymorphisms are a consistent feature and can be assessed readily from normal tissues (blood cells and skin) using standard techniques. Recent studies, including research from our laboratory, have suggested that one polymorphism of the EGFR gene, a highly polymorphic dinucleotide CA repeat in intron 1, related to the transcriptional activity of the EGFR gene, can predict the outcome of EGFR-TKI treatment in advanced NSCLC patients.[12–15] Patients with shorter (CA)n repeats ($n \leqslant 16$) have improved clinical response and longer survival than those with longer (CA)n repeats ($n > 16$). These studies confirm that inherited genetic factors, such as genetic polymorphisms, can influence the results of EGFR-TKI treatment.

CYP1A1 is a highly inducible member of the cytochrome P450 family, which is expressed in human lung.[16,17] It is involved in the metabolism of oestradiol through catalysing C-2 hydroxylation, and may influence EGFR protein expression.[18] It has been demonstrated that CYP1A1 not only metabolises oestrogens, but also plays a central role in the activation of major classes of tobacco carcinogens, including polyaromatic hydrocarbons, and produces many reactive intermediates (e.g. catechol oestrogen, which can cause gene mutations). To date, four polymorphic forms of CYP1A1 (CYP1A1*2A–*2D) have been identified and their genotypes appear to exhibit population frequencies that vary by ethnicity; significantly higher frequencies of CYP1A1*2A and CYP1A1*2C alleles have been reported amongst Asians versus Caucasians and African-Americans.[19,20] The CYP1A1*2A allele has a T → C mutation in the 3′ non-coding region (3801 T → C), which has been associated with elevated enzyme activity.[21] An A → G transition in exon 7 creates the second allelic variant (*2C), leading to the amino acid substitution of Val for Ile in the haeme-binding region, resulting in an increase in microsomal enzyme activity.[22] There is increasing evidence that individuals with homozygous CYP1A1 (T/T and A/A) genotypes are at higher risk of smoking-associated lung cancer, as seen in Asian population studies where these 'at risk' allele frequencies are 8–18 times higher than in Caucasians.[23]

As either the EGFR intron 1 or the CYP1A1 gene polymorphisms may affect EGFR expression and function, which determines the outcome of EGFR-TKI treatment, we hypothesised that genetic variations in these two genes might affect clinical outcome in NSCLC patients treated with EGFR-TKIs. Thus, in our study, we examine whether these polymorphisms in the EGFR and CYP1A1 gene can serve as molecular markers predicting EGFR-TKI response and overall survival in NSCLC patients.

2. Patients and methods

2.1. Eligible subjects

In total, 115 patients with histological confirmation (25 confirmed by lymph node excision, 16 by bronchoscopy biopsy, and 74 by CT guided biopsy) of stage IIIb or IV NSCLC who received gefitinib ($n = 70$) or erlotinib ($n = 45$) through the Expanded Assess Program (EAP)[14] were enrolled. Patients were enrolled from June 2002 to September 2006 at the Guangdong General Hospital. All enrolled patients received at least one course of standard systemic chemotherapy prior to the EAP. The final follow-up was in July 2010.

The study was approved by the Ethics Committee. Written informed consent was obtained from all patients.

2.2. Clinical evaluation and response criteria

Enrolled patients received gefitinib (250 mg) or erlotinib (150 mg) daily. Tumours were evaluated by CT scans performed pre-treatment and 4 and 10 weeks after treatment initiation. If patients were confirmed to have a tumour response or stable disease, they continued on EGFR-TKI and had repeated tumour evaluations every 12 weeks until tumour progression. A treatment response was defined according to the RECIST Criteria.[24]

2.3. DNA extraction and genotyping

2.3.1. DNA extraction

Blood was collected from each patient prior to starting EGFR-TKI treatment. Genomic DNA was extracted from white blood cells using a Universal Genomic DNA Extraction Kit (ver. 3.0; Takara, Dalian). Briefly, 200 μL of blood was mixed with 500 μL Solution A and 1 μL RNase A1 in the collection tube, and vibrated for 15 s, and placed on ice for 5 min. To the mixture, 400 μL Solution B was added, vibrated for 15 s and followed by the addition of 1 mL Solution C. The resulting mixture was centrifuged (12,000 rpm, 2 min). The upper organic phase was discarded and the final DNA extraction was performed following the manufacturer's protocol.

The presence of the EGFR R497K, CYP1A1*2A and CYP1A1*2C polymorphisms was tested using a PCR-restriction fragment length polymorphism (RFLP) method, as described previously.[25,26] Forward and reverse primers, restriction enzymes, and annealing temperatures for specific gene polymorphisms are listed in Table 1. Briefly, forward and reverse primers were used for PCR amplification, the PCR product was digested with restriction enzymes (New England Biolabs, Beverly, MA, USA), and alleles were separated on 4% (EGFR R497K) or 2% (CYP1A1*2A and *2C) Nusieve ethidium bromide-stained agarose gels. For the EGFR R497K polymorphism, samples from individuals with homozygous wild-type

Table 1 – Primers, annealing temperatures and restriction enzymes.

Gene	Forward primer (5′ → 3′)	Reverse primer (5′ → 3′)	Annealing temperature (°C)	Restriction enzymes
EGFR intron 1	GGATCGCGGGACTCTTGA	GTGGGTTTATGGTCGGTA	55	NA
EGFR R497K	TGCTGTGACCCACTCTGTCT	CCAGAAGGTTGCACTTGTCC	59	BstN1
CYP1A1*2A	CAGTGAAGAGGTGTAGCCGCT	TAGGAGTCTTGTCTCATGCCT	65	MspI
CYP1A1*2C	GTAGACAGAGTCTAGGCCTCA	GAAGTGTATCGGTGAGACCA (Ile) GAAGTGTATCGGTGAGACCG (Val)	65	NA

EGFR, epidermal growth factor receptor; CYP, cytochrome P450.

alleles produced one undigested 102-bp band, samples from heterozygous individuals gave bands positioned at 102, 54 and 48 bp, and if a patient had a homozygous mutation, only bands at 54- and 48-bp were evident. For the CYP1A1*2A polymorphism, genotyping with homozygous wild-type individuals was characterised by a 340-bp fragment; genotyping with heterozygous individuals were characterised by 340-, 200- and 140-bp fragments, and a homozygous mutation was characterised by 200- and 140-bp fragments. For CYP1A1*2C polymorphism, samples that were homozygous for the Ile form yielded the 210-bp product only when primer Ile was used; samples homozygous for the Val form yielded the 210-bp product only when primer Val was used, and if the 210-bp product was obtained with both Ile and Val primer, the sample was deemed to be heterozygous.

2.3.2. *CA repeats in intron 1*

PCR amplification was performed in a 50 μL total reaction volume containing 100 ng of genomic DNA, 1.5 mM $MgCl_2$, 0.2 mM each dNTP, 1× PCR amplification buffer, 1.5 units of Taq polymerase (Promega) and 0.4 μL of each primer. Primer sequences specific for this microsatellite marker are listed in Table 1. Amplification conditions were 35 cycles of 95 °C for 1 min, 55 °C for 45 s and 72 °C for 1 min. CA dinucleotide repeats of the PCR products were directly sequenced with the ABI3100 genetic analyzer, using the BigDye Terminator kit (ver. 3.1; Applied Biosystems). Then, 40 ng of genomic DNA was amplified by PCR with 5′-FAM – labelled forward EGFR primer (5′-GGA-TCG-CGG-GAC-TCT-TGA-3′) and unlabelled reverse EGFR primer (5′-GTG-GGT-TTA-TGG-TCG-GTA-3′). PCR conditions were as described earlier in this section. After PCR, 2 μL of the product was denatured in a 10-μL mix of formamide and Genescan 500 LIZ molecular weight standard (Applied Biosystems) at 95 °C for 5 min, separated with an Applied Biosystems Prism Genetic Analyzer with POP4 polymer and fragment lengths were determined. Genotypes were resolved on the Applied Biosystems ABI3100 genetic analyzer (GeneMapper Software, ver. 3.5) for allele length determination.

2.4. *EGFR mutation analysis and IHC assay*

Mutational analysis was performed as described previously.[8] Briefly, DNA was extracted from frozen tumour tissues with the Trizol reagent (Invitrogen, USA), according to the manufacturer's protocol. The kinase domains of EGFR 18–21 exons were amplified with four pairs of primers (exon 18: 5′-AGG-GCT-GAG-GTG-ACC-CTT-GT-3′ and 5′-TCC-CCA-CCA-GAC-CAT-GAG-AG-3′; exon 19: 5′-ACC-ATC-TCA-CAA-TTG-CCA-GTT-AAC-3′ and 5′-GAG-GTT-CAG-AGC-CAT-GGA-CC-3′; exon 20: 5′-ACT-GAC-GTG-CCT-CTC-CCT-C-3′ and 5′-CCC-GTA-TCT-CCT-TCC-CTG-3′; exon 21: 5′-TCA-CAG-CAG-GGT-CTT-CTC-TGT-TT-3′ and 5′-ATG-CTG-GCT-GAC-CTA-AAG-CC-3′). Amplification conditions were 40 cycles of 94 °C for 35 s, 56 °C for 60 s, and 72 °C for 1 min. PCR amplicons were analysed by direct sequencing. Reactions were electrophoresed on an ABI 3100 genetic analyzer and sequence variations were determined using the SeqScape software (Applied Biosystems, USA).

Expression of the EGFR protein in tumour tissues was detected by IHC using Dako EGFR PharmDx kits (Dakocytomation, Denmark). Briefly, 4 μm tissue sections were deparaffinised in xylene and hydrated through graded alcohols. EGFR status was assessed using a 1/40 diluted mAb anti-EGFR at room temperature for 1 h, and endogenous peroxidase activity was blocked by 30 min treatment with 3% hydrogen peroxide in absolute methanol at room temperature. EGFR immunostaining was scored as 0 (0–10% of considered cells stained), 1+(11–25%), 2+(26–50%) and 3+(>50%). Stained slices were read by the pathology department in our hospital.

2.5. *Statistical analyses*

Associations of each polymorphism and clinical-baseline characteristics (age, gender, TNM stage, histology, smoking status, performance, side-effects) with the response to EGFR-TKIs were summarised using contingency Tables and non-parametric tests (Mann–Whitney or Kruskal–Wallis, depending on the classification of genotypes). Odds ratios and 95% confidence intervals (CI) were calculated by multivariate logistic regression analysis, adjusting for known or potential clinical predictor factors, such as gender, TNM stage, histology, smoking index and side-effects. Overall survival (OS) was defined as the time from the beginning of EGFR-TKI to death from any cause, and was estimated using the Kaplan–Meier method and the log-rank test. Median follow-up time was computed with censored observations only. Median survival time (MST) was calculated. Statistical analyses were performed using the SPSS software. All calculated p values were two-sided.

3. Results

3.1. *Patient characteristics*

In total, 115 patients were enrolled in this study, 65 (57%) men and 50 (43%) women, with a median age of 57 years. Patient characteristics are summarised in Table 2. Follow-up times

ranged from 47 to 83 months with a median follow-up time of 54 months for surviving patients. The data cut-off was July 20, 2010. A partial response to EGFR-TKIs was observed in 42.6% of patients (49/115), stable disease in 19.1% (22/115), and progressive disease in 38.3% (44/115). Treatment outcomes to EGFR-TKIs were correlated with clinicopathological characteristics including age, gender, tumour TNM staging, histology, smoking index, performance status and skin toxicity (Table 2). There was no statistically significant association except those of the smoking index and histology status ($p = 0.003$, $p = 0.009$, respectively; Pearson chi-squared test).

3.2. Relationship between EGFR mutations, EGFR expression, and gene polymorphisms

For the subgroup of 36 patients whose tumour tissues were available, EGFR mutation analysis showed that 19 patients had a somatic mutation (12 in-frame deletions in exon 19 (del E746-A750), 7 substitution mutations in exon 21 (L858R)) and the remaining 17 patients were wild-type. These limited data showed no correlation between EGFR somatic mutation and intron 1 (CA)n or R497K polymorphisms of EGFR ($p = 0.194$, $p = 0.977$, respectively; Pearson chi-squared test; Table 3). However, the *2A, *2C polymorphism of the CYP1A1 gene showed a correlation with EGFR mutations ($p = 0.037$, $p = 0.021$, respectively; Pearson chi-squared test; Table 3). A further subgroup of 17 patients in the cohort (including 3 patients who did not have an EGFR mutation test) had tumour tissues available for IHC; these data demonstrated no association between EGFR protein expression and the genomic polymorphisms except for the EGFR intron 1 (CA)n repeat polymorphism ($p = 0.035$, Fisher's exact test; Table 3).

3.3. Gene polymorphisms and clinical response to EGFR-TKIs

The CYP1A1*2A polymorphism showed a significant association with the clinical response to EGFR-TKIs. Patients with the homozygous T/T genotype had better clinical responses than patients with heterozygous T/C and C/C ($p = 0.011$; Table 4). Of the 41 patients with T/T genotypes, 23 (56%) showed a partial response to EGFR-TKIs, nine (22%) had stable disease and only nine (22%) were observed to have progressive disease after EGFR-TKI therapy. Additionally, the EGFR intron 1 polymorphism showed an association with clinical response to EGFR-TKIs. Patients with shorter repeats

Table 2 – Patient baseline characteristics and correlations with clinical outcome to EGFR-TKIs ($n = 115$).

Characteristics*	Number of patients	Patients responding to EGFR-TKI			p Value[a]	Survival time		p Value[b]
		PR	SD	PD		Median months (95% CI)	Relative risk (95% CI)	
Gender					0.278			0.375
Female	50	24	11	15		18 (10.9–26.8)	1 (reference)	
Male	65	25	11	29		9 (4.4–13.8)	1.19 (0.81–1.75)	
Age					0.500			0.156
⩽57	60	27	9	24		11 (6.7–15.5)	(reference)	
>57	55	22	13	20		18 (12.3–24.6)	1.32 (0.90–1.94)	
Smoking index					0.003			0.000
Non-smoker	65	32	15	18		20 (13.9–26.7)	1 (reference)	
⩽400	23	13	1	9		15 (3.6–28.2)	0.38 (0.24–0.61)	
>400	27	4	6	17		5 (2.5–7.4)	0.42 (0.23–0.75)	
Histological status					0.009			0.000
ADC	96	46	19	31		17 (12.5–22.1)	1 (reference)	
Others	19	3	3	13		4 (2.9–5.1)	0.39 (0.23–0.65)	
TNM stage					0.692			0.001
IV	96	43	18	35		16 (11.2–20.8)	1 (reference)	
Others	17	6	3	8		6 (3.0–9.0)	2.40 (1.40–4.12)	
Unspecified	2							
Performance status					0.405			0.000
0–1	99	44	20	35		18 (13.1–23.8)	1 (reference)	
2–3	15	5	2	8		4 (0.2–8.2)	0.24 (0.13–0.43)	
No record	1							
Skin toxicity					0.159			0.980
0–1	63	24	16	23		15 (8.5–21.5)	1 (reference)	
2–3	43	24	6	13		16 (10.5–21.5)	0.99 (0.67–1.50)	
No record	9							

Abbreviations: ADC, adenocarcinoma; PR, partial response; SD, stable disease; PD, progressive disease.

*Smoking index was evaluated by the Brinkmann index; performance status was commented by Ecog–Who–Zubord score; skin toxicity was evaluated according to NCI CTCAE, ver.3.

[a] Pearson chi-squared test.

[b] log-rank test.

Table 3 – Associations between EGF receptor mutation and IHC and EGFR and CYP1A1 polymorphisms.

Polymorphisms	Number of patients[#]	EGFR expression tumour tissues (%)		p Value[a]	Number of patients	Somatic mutation of EGFR (%)		p Value[b]
		Negative (0 or +)	Positive (++ or +++)			Wild type	Mutated type	
EGFR intron 1(CA)n				0.035				0.194
Any allele ⩽16CA	11	4 (36.4)	7 (63.6)		21	8 (38.1)	13 (61.9)	
Both alleles >16CA	6	6 (100.0)	0 (0.0)		15	9 (60.0)	6 (40.0)	
EGFR R497K				0.838				0.977
A/A	6	3 (50.0)	3 (50.0)		11	5 (45.5)	6 (54.5)	
A/L	7	4 (57.1)	3 (42.9)		15	7 (46.7)	8 (53.5)	
L/L	4	3 (75.0)	1 (25.0)		10	5 (50.0)	5 (50.0)	
CYP1A1*2A				0.35				0.037
T/T	7	3 (42.9)	4 (57.1)		15	4 (26.7)	11 (73.3)	
T/C + C/C	10	7 (70.0)	3 (30.0)		21	13 (61.9)	8 (38.1)	
CYP1A1*2C				1.000				0.021
A/A	8	5 (62.5)	3 (37.5)		11	2 (18.2)	9 (81.8)	
A/G + G/G	9	5 (55.6)	4 (44.4)		25	15 (60.0)	10 (40.0)	

[a] Fisher's exact test.
[b] Pearson chi-squared test.
[#] Three patients did not have EGFR mutation tests.

(⩽16 CA repeats) had significantly improved clinical responses versus patients with longer repeats (>16 CA repeats; $p = 0.046$; Table 4). Of the 66 patients with shorter repeats, 50% (33/66) showed a partial response to EGFR-TKIs, whereas 47% (23/49) of patients with longer repeats showed progressive disease despite EGFR-TKI therapy. Multivariate analysis confirmed that both polymorphisms were independent predictive factors ($p = 0.028$, $p = 0.007$, respectively). For the subgroup of 36 patients whose tumour tissues were available, we also detected the EGFR mutation status and the clinical response to EGFR-TKI. These data showed that patients with EGFR mutations had an improved response versus those with

Table 4 – Correlation between genomic polymorphisms and clinical outcomes to EGFR-TKIs.

Polymorphisms	Number of patients	Patients responding to EGFR-TKI (%)			p[a] Value	Survival		p[b] Value
		PR	SD	PD		Median months (95% CI)	Relative risk (95% CI)	
CYP1A1*2A[#]					0.011			0.001
T/T	41	23 (56.0)	9 (22.0)	9 (22.0)		23.5 (11.7–35.2)	1 (reference)	
T/C + C/C	74	26 (35.1)	13 (17.6)	35 (47.3)		10.2 (3.4–17.1)	0.48 (0.31–0.73)	
CYP1A1*2C					0.629			0.172
A/A	28	13 (46.4)	7 (25.0)	8 (28.6)		18.7 (14.9–22.4)	1 (reference)	
A/G + G/G	87	36 (41.4)	15 (17.2)	36 (41.4)		11.1 (4.7–17.4)	0.74 (0.47–1.16)	
EGFR intron 1 (CA)n					0.046			0.293
Any allele ⩽16CA	66	33 (50.0)	12 (18.2)	21 (31.8)		15.9 (9.4–22.4)	1 (reference)	
Both alleles >16CA	49	16 (32.7)	10 (20.4)	23 (46.9)		10.7 (4.7–16.8)	0.81 (0.55–1.19)	
EGFR R497K					0.573			0.299
A/A	43	19 (44.2)	8 (18.6)	16 (37.2)		17.4 (3.4–31.5)	1 (reference)	
A/L	47	20 (42.6)	9 (19.1)	18 (38.3)		14.3 (7.0–21.6)	0.81 (0.49–1.34)	
L/L	25	10 (40.0)	5 (20.0)	10 (40.0)		12.3 (1.1–23.4)	0.71 (0.43–1.17)	
CYP1A1*2A+EGFR intron 1 (CA)n					0.002			0.003
T/T+ any allele ⩽16CA	22	17 (77.3)	4 (18.2)	1 (4.5)		27.8 (22.4–33.2)	1 (reference)	
T/C + C/C + both alleles >16CA	93	32 (34.4)	18 (19.4)	43 (46.2)		10.8 (5.3–16.4)	0.46 (0.28–0.77)	

[a] Adjusted by gender, TNM stage, histology status, smoking index and side-effects.
[b] Log-rank test.
[#] Only nine cases had the C/C genotype of CYP1A1*2A, and were analysed together with the T/C genotype.

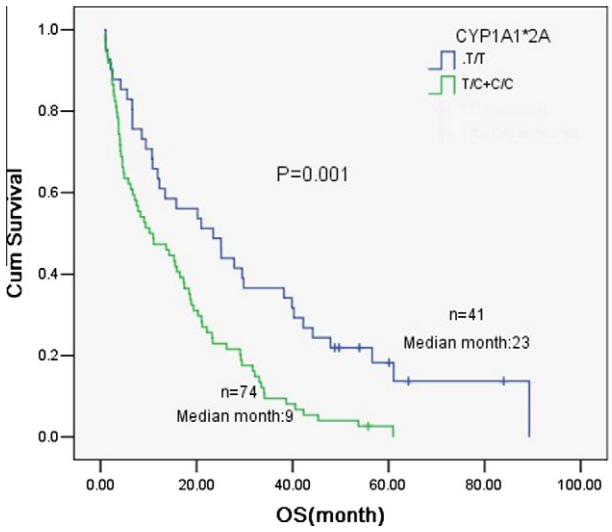

Fig. 1 – Kaplan–Meier curve of survival time by CYP1A1*2A polymorphism.

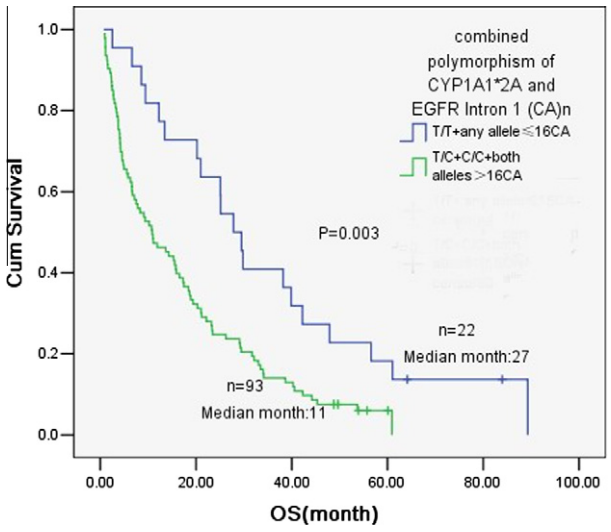

Fig. 2 – Combined effects of EGFR intron1 (CA)n and CYP1A1*2A polymorphisms on overall survival.

wild-type EGFR ($p = 0.001$), consistent with previous research.[6] Nevertheless, no link was identified between the EGFR R497K or CYP1A1*2C polymorphism and clinical response to EGFR-TKIs ($p = 0.573$, $p = 0.629$, respectively; Table 4).

3.4. *Gene polymorphisms and overall survival*

At time of analysis, sadly 108 (93.9%) patients had died; only seven (6.1%) patients were alive. The median survival time was 13 months (95% CI = 8–18 months). Univariate survival analysis (log-rank test) demonstrated that only the CYP1A1*2A polymorphism showed a significant correlation with overall survival ($p = 0.001$, log-rank test; Table 4, Fig. 1). Cox regression analysis confirmed that the CYP1A1*2A polymorphism was a strong independent prognosis factor ($p < 0.001$; Table 5). However, no significant difference was found amongst EGFR intron 1 (CA)n, R497K, and CYP1A1*2C polymorphisms and patient survival ($p = 0.293$, $p = 0.299$, $p = 0.172$, respectively; log-rank test, Table 4).

Table 5 – Multivariate analysis on overall survival (Cox regression model).

Prognosis factors*	Relative risk (RR)	95% CI (%)	*p* Value
Histological status	3.184	1.853–5.474	0.000
Performance status	4.713	2.596–8.558	0.000
CYP1A1*2A	2.364	1.535–3.639	0.000
Combined polymorphisms of EGFR intron 1 (CA)n and CYP1A1*2A	1.992	1.182–3.358	0.010

* Smoking index and TNM stage were included in the Cox regression model because the log-rank test showed the factors had statistical significance with regard to overall survival. The CYP1A1*2A polymorphism, the combined polymorphisms of EGFR intron 1 (CA)n and CYP1A1*2A were analysed alone with the other four base line factors (TNM stage, smoking index, histology status, performance status).

3.5. *Combined genotype effects on clinical outcome*

To combine the analysis with the EGFR intron 1 genotype, we first divided CYP1A1*2A genotypes into two groups: patients with only the T/T genotype and patients with T/C and C/C allele genotypes. Results of the analysis in combination with EGFR intron 1 (CA)n repeats and CYP1A1*2A genotypes showed a striking association with patient response to EGFR-TKIs and overall survival. Of the 22 patients with any allele ⩽16 CA repeats of the EGFR intron 1 genotype and T/T homozygous for the CYP1A1 genotype, 17 (77.3%) showed a partial response to EGFR-TKIs, four (18.2%) had stable disease, and only one (4.5%) had progressive disease. In contrast, 43 of 93 patients (46.2%) who were both T/C heterozygous or C/C homozygous for the CYP1A1 genotype and both alleles >16 CA repeats of EGFR intron 1 genotypes were observed to have progressive disease despite EGFR-TKIs, 18 (19.4%) had stable disease, and 32 (34.4%) had a partial response to EGFR-TKIs ($p = 0.002$; Table 4). The former group of patients had a median survival time of 27 months (95% CI = 18–35 months), whereas the latter group of patients survived only 11 months (95% CI = 7–15 months; $p = 0.003$, log-rank test; Table 4, Fig. 2).

4. Discussion

Mechanisms that cause patients to have differing levels of sensitivity to EGFR-TKIs remain largely unknown. Cancer is a genetic disease, and EGFR is a validated anticancer target the successful exploitation of which has added novel agents, such as gefitinib, erlotinib and cetuximab to the current treatment options.[27,28] Recent studies in patients with NSCLC have highlighted the importance of a patient's genetic make-up in determining the response to EGFR-TKIs.[7,8] Subsets of patients have been shown to benefit the most from these therapies, and although the differential traits have yet to be completely defined, they are mostly genetic.[28]

In the present study, we evaluated the effects of four polymorphisms in the EGFR and CYP1A1 genes on the clinical outcome to EGFR-TKI therapy in NSCLC patients. The results demonstrated that two gene polymorphisms, EGFR intron 1 (CA)n repeats and CYP1A1*2A, and the joint effects of both of these genetic polymorphisms, were associated with clinical outcome to EGFR-TKI therapy in advanced NSCLC patients.

The EGFR gene contains a highly polymorphic sequence in intron 1, consisting of a variable number of CA dinucleotide repeats, ranging from 9 to 21. This sequence affects the efficiency of gene transcription such that subjects or cell lines with a larger number of CA repeats have lower levels of mRNA and protein expression.[29,30] For example, transcription of EGFR is inhibited by approximately 80% in alleles with 21 CA repeats. Recent studies have shown that in patients with NSCLC, an increased EGFR gene copy number and high EGFR protein expression are significantly associated with improved response and survival after gefitinib therapy.[31,32] Moreover, Taron et al. reported that EGFR intron 1 (CA)n polymorphisms tended to be associated with EGFR mutations.[33] Our results confirm that EGFR intron 1 (CA)n polymorphism is associated with the clinical response to EGFR-TKI of NSCLC patients ($p = 0.046$; Table 4). However, our long-term follow-up data demonstrated a negative result between the EGFR intron 1 (CA)n polymorphism and patient overall survival ($p = 0.293$; Table 4). Our results showed that the EGFR intron 1 (CA)n polymorphism was linked to EGFR protein expression in tumour tissues, and this may be part of the reason why the EGFR intron 1 (CA)n polymorphism is associated with improved clinical response to EGFR-TKI therapy. However, the level of EGFR protein expression may have an adverse effect on overall survival (OS) in NSCLC patients. Selvaggi et al. reported that in 130 NSCLC patients undergoing an operation, those with EGFR overexpression had shorter survival compared with patients with lower EGFR expression.[34] This indicates that the EGFR intron 1 (CA)n polymorphism may serve as a predictive factor, but not a prognostic factor, for patients being treated with EGFR-TKI therapy. We found no correlation between the intron 1 polymorphisms and EGFR somatic mutations.

CYP1A1 is a phase-one enzyme that metabolises oestrogen and a variety of environmental carcinogens, including polyaromatic hydrocarbons, and produces many reactive intermediates, such as catechol oestrogen. This enzyme catalyses the C2-, C6- and C15- hydroxylation of oestradiol. CYP1A1 is also involved in the metabolism of xenobiotics. Studies have confirmed that CYP1A1 polymorphisms are significantly correlated with the risk of lung, breast, head and neck and oesophageal cancer. CYP1A1*2A polymorphism is also a prognostic factor for NSCLC patients. Goto et al. reported that patients with heterozygous alleles had shortened survival compared with those with homozygous alleles, especially in the subgroup of patients with a history of smoking.[35] However, there is no reported analysis of the relationship between CYP1A1 gene polymorphisms and the treatment of cancer. Our study is the first to report that the CYP1A1*2A polymorphism correlates with the response to EGFR-TKI of NSCLC. We found that T/T homozygous patients had an improved response to EGFR-TKI versus those with heterozygous T/C and homozygous C/C alleles ($p = 0.011$; Table 4). The Cox regression model also showed that T/T homozygous patients gave a strong independent prognosis factor for NSCLC patients treated with EGFR-TKIs ($p < 0.001$; Table 5). When we combined the EGFR intron 1 (CA)n repeats and CYP1A1*2A polymorphism analysis, we found that the joint effects of the polymorphisms were strongly associated with the response to EGFR-TKIs and overall survival of NSCLC patients treated with EGFR-TKIs ($p = 0.003$, log rank test). The CYP1A1*2C polymorphism showed no association with the response to EGFR-TKIs or the overall survival of the NSCLC patients ($p = 0.629$, Pearson chi-squared test; $p = 0.172$, log rank test, respectively), and interestingly, our small sample size showed both SNPs of the CYP1A1*2A, CYP1A1*2C gene to correlate with EGFR somatic mutations ($p = 0.037$, $p = 0.021$; respectively).

As discussed, CYP1A1 is a major enzyme involved in oestrogen hydroxylation. Recent studies have shown that oestrogen metabolites, such as catechol oestrogen, 4-hydroxyestrone and 16a-hydroxyestrone, can bind to DNA, creating adducts and causing gene mutations,[36] which may explain the association between both SNPs of the CYP1A1 gene and the EGFR mutations. However, several reports suggest that oestrogen may be involved in the etiology of lung cancer,[37] and treatment of lung cancer with an anti-oestrogenic drug could give rise to a population of cells with increased levels of EGFR, because the EGFR and ER pathway have a functional linkage.[38] Stabile et al. reported that the presence of oestrogen alone is sufficient to decrease EGFR expression in the A549 lung cancer cell line, showing that oestrogen can modulate EGFR levels.[39] These findings may help to explain how the synergistic effects of the CYP1A1*2A and the EGFR intron 1 (CA)n repeat polymorphisms influence patient clinical outcomes to EGFR-TKIs. One possible mechanism may be that the T/T genotype of CYP1A1 increases oestrogen metabolism, decreasing oestrogen levels, thereby increasing the levels of EGFR expression; the shorter intron 1 CA repeats (any allele $\leqslant$16 CA repeats) of the EGFR gene also increased EGFR mRNA transcription and EGFR expression. Thus, patients with both the T/T allele of CYP1A1 and shorter intron 1 CA repeats (any allele $\leqslant$16 CA repeats) of the EGFR gene may have enhanced effects and show improved clinical outcomes to EGFR-TKIs. However, further studies are needed to investigate the exact mechanism of these clinical findings.

To conclude, our study demonstrated that the CYP1A1*2A polymorphism is a potential surrogate predictive and prognostic marker for NSCLC patients treated with EGFR-TKIs. More importantly, we found that combined analysis of two polymorphisms showing a synergistic effect on the response to EGFR-TKI had a significant association with NSCLC patient clinical outcome to EGFR-TKIs. If this is confirmed by further studies with larger numbers of patients, such combined analysis may offer a useful predictor of clinical outcomes to EGFR tyrosine kinase inhibitor therapy.

Conflict of interest statement

None declared.

Acknowledgements

This study was supported by grants from the National Natural Science Foundation of China (to Qiang Nie, No. 30871126), the National Science Foundation of Guangdong Province (to Qiang Nie, No. 7300046), and the Foundation of Guangdong Science and Technology Department (to Y. Wu, No. 2006B60101010).

REFERENCES

1. Hung RJ, Boffetta P, Brockmoller J, et al. CYP1A1 and GSTM1 genetic polymorphisms and lung cancer risk in Caucasian non-smokers: a pooled analysis. *Carcinogenesis* 2003;**24**:875–82.
2. Cohen MH, Williams GA, Sridhara R, et al. United States Food and Drug Administration Drug Approval summary: Gefitinib (ZD1839; Iressa) tablets. *Clin Cancer Res* 2004;**10**:1212–8.
3. Miller VA, Kris MG, Shah N, et al. Bronchioloalveolar pathologic subtype and smoking history predict sensitivity to gefitinib in advanced non-small-cell lung cancer. *J Clin Oncol* 2004;**22**:1103–9.
4. Ciardiello F, Tortora G. EGFR antagonists in cancer treatment. *N Engl J Med* 2008;**358**:1160–74.
5. Pao W, Miller V, Zakowski M, et al. EGF receptor gene mutations are common in lung cancers from 'never smokers' and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci USA* 2004;**101**:13306–11.
6. Mok TS, Wu YL, Thongprasert S, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 2009;**361**:947–57.
7. Mitsudomi T, Kosaka T, Endoh H, et al. Mutations of the epidermal growth factor receptor gene predict prolonged survival after gefitinib treatment in patients with non-small-cell lung cancer with postoperative recurrence. *J Clin Oncol* 2005;**23**:2513–20.
8. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 2004;**350**:2129–39.
9. Marchetti A, Martella C, Felicioni L, et al. EGFR mutations in non-small-cell lung cancer: analysis of a large series of cases and development of a rapid and sensitive method for diagnostic screening with potential implications on pharmacologic treatment. *J Clin Oncol* 2005;**23**:857–65.
10. Uramoto H, Sugio K, Oyama T, et al. Epidermal growth factor receptor mutations are associated with gefitinib sensitivity in non-small cell lung cancer in Japanese. *Lung cancer* 2006;**51**:71–7.
11. Tiseo M, Rossi G, Capelletti M, et al. Predictors of gefitinib outcomes in advanced non-small cell lung cancer (NSCLC): study of a comprehensive panel of molecular markers. *Lung cancer* 2010;**67**:355–60.
12. Tiseo M, Capelletti M, De Palma G, et al. Epidermal growth factor receptor intron-1 polymorphism predicts gefitinib outcome in advanced non-small cell lung cancer. *J Thorac Oncol* 2008;**3**:1104–11.
13. Ichihara S, Toyooka S, Fujiwara Y, et al. The impact of epidermal growth factor receptor gene status on gefitinib-treated Japanese patients with non-small-cell lung cancer. *Int J Cancer* 2007;**120**:1239–47.
14. Liu G, Gurubhagavatula S, Zhou W, et al. Epidermal growth factor receptor polymorphisms and clinical outcomes in non-small-cell lung cancer patients treated with gefitinib. *Pharmacogenomics J* 2008;**8**:129–38.
15. Nie Q, Wang Z, Zhang GC, et al. The epidermal growth factor receptor intron1 (CA) n microsatellite polymorphism is a potential predictor of treatment outcome in patients with advanced lung cancer treated with Gefitinib. *Eur J Pharmacol* 2007;**570**:175–81.
16. Saarikoski ST, Husgafvel-Pursiainen K, Hirvonen A, et al. Localization of CYP1A1 mRNA in human lung by in situ hybridization: comparison with immunohistochemical findings. *Int J Cancer* 1998;**77**:33–9.
17. Hukkanen J, Lassila A, Paivarinta K, et al. Induction and regulation of xenobiotic-metabolizing cytochrome P450s in the human A549 lung adenocarcinoma cell line. *Am J Respir Cell Mol Biol* 2000;**22**:360–6.
18. Mollerup S, Berge G, Baera R, et al. Sex differences in risk of lung cancer: Expression of genes in the PAH bioactivation pathway in relation to smoking and bulky DNA adducts. *Int J Cancer* 2006;**119**:741–4.
19. Garte S. The role of ethnicity in cancer susceptibility gene polymorphisms: the example of CYP1A1. *Carcinogenesis* 1998;**19**:1329–32.
20. Cascorbi I, Brockmoller J, Roots I. A C4887A polymorphism in exon 7 of human CYP1A1: population frequency, mutation linkages, and impact on lung cancer susceptibility. *Cancer Res* 1996;**56**:4965–9.
21. Kiyohara C, Hirohata T, Inutsuka S. The relationship between aryl hydrocarbon hydroxylase and polymorphisms of the CYP1A1 gene. *Jpn J Cancer Res* 1996;**87**:18–24.
22. Kiyohara C, Nakanishi Y, Inutsuka S, et al. The relationship between CYP1A1 aryl hydrocarbon hydroxylase activity and lung cancer in a Japanese population. *Pharmacogenetics* 1998;**8**:315–23.
23. Bartsch H, Nair U, Risch A, et al. Genetic polymorphism of CYP genes, alone or in combination, as a risk modifier of tobacco-related cancers. *Cancer Epidemiol Biomarkers Prev* 2000;**9**:3–28.
24. Therasse P, Arbuck SG, Eisenhauer EA. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst* 2000;**92**:205–16.
25. Zhang W, Park DJ, Lu B, et al. Epidermal growth factor receptor gene polymorphisms predict pelvic recurrence in patients with rectal cancer treated with chemoradiation. *Clin Cancer Res* 2005;**11**:600–5.
26. Noda N, Matsuzoe D, Konno T, et al. Risk for K-ras gene mutations in smoking-induced lung cancer is associated with cytochrome P4501A1 and glutathione S-transferase micro1 polymorphisms. *Oncol Rep* 2004;**12**:773–9.
27. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med* 2004;**10**:789–99.
28. Jimeno A, Hidalgo M. Pharmacogenomics of epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors. *Biochim Biophys Acta*, 2006.
29. Araujo A, Ribeiro R, Azevedo I, et al. Genetic polymorphisms of the epidermal growth factor and related receptor in non-small cell lung cancer – a review of the literature. *Oncologist* 2007;**12**:201–10.
30. Nomura M, Shigematsu H, Li L, et al. Polymorphisms, mutations, and amplification of the EGFR gene in non-small cell lung cancers. *PLoS Med* 2007;**4**:3125–210.
31. Cappuzzo F, Hirsch FR, Rossi E, et al. Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer. *J Natl Cancer Inst* 2005;**97**:643–55.
32. Hirsch FR, Varella-Garcia M, McCoy J, et al. Increased epidermal growth factor receptor gene copy number detected by fluorescence in situ hybridization associates with increased sensitivity to gefitinib in patients with

bronchioloalveolar carcinoma subtypes: a Southwest Oncology Group Study. *J Clin Oncol* 2005;**23**:6838–45.
33. Taron M, Ichinose Y, Rosell R, et al. Activating mutations in the tyrosine kinase domain of the epidermal growth factor receptor are associated with improved survival in gefitinib-treated chemorefractory lung adenocarcinomas. *Clin Cancer Res* 2005;**11**:5878–85.
34. Selvaggi G, Novello S, Torri V, et al. Epidermal growth factor receptor overexpression correlates with a poor prognosis in completely resected non-small-cell lung cancer. *Ann Oncol* 2004;**15**:28–32.
35. Goto I, Yoneda S, Yamamoto M, et al. Prognostic significance of germ line polymorphisms of the CYP1A1 and glutathione S-transferase genes in patients with non-small cell lung cancer. *Cancer Res* 1996;**56**:3725–30.
36. Service RF. New role for estrogen in cancer? *Science* 1998;**279**:1631–3.
37. Patel JD, Bach PB, Kris MG. Lung cancer in US women: a contemporary epidemic. *JAMA* 2004;**291**:1763–8.
38. Levin ER. Bidirectional signaling between the estrogen receptor and the epidermal growth factor receptor. *Mol Endocrinol* 2003;**17**:309–17.
39. Stabile LP, Lyker JS, Gubish CT, et al. Combined targeting of the estrogen receptor and the epidermal growth factor receptor in non-small cell lung cancer shows enhanced antiproliferative effects. *Cancer Res* 2005;**65**:1459–70.