



OPEN

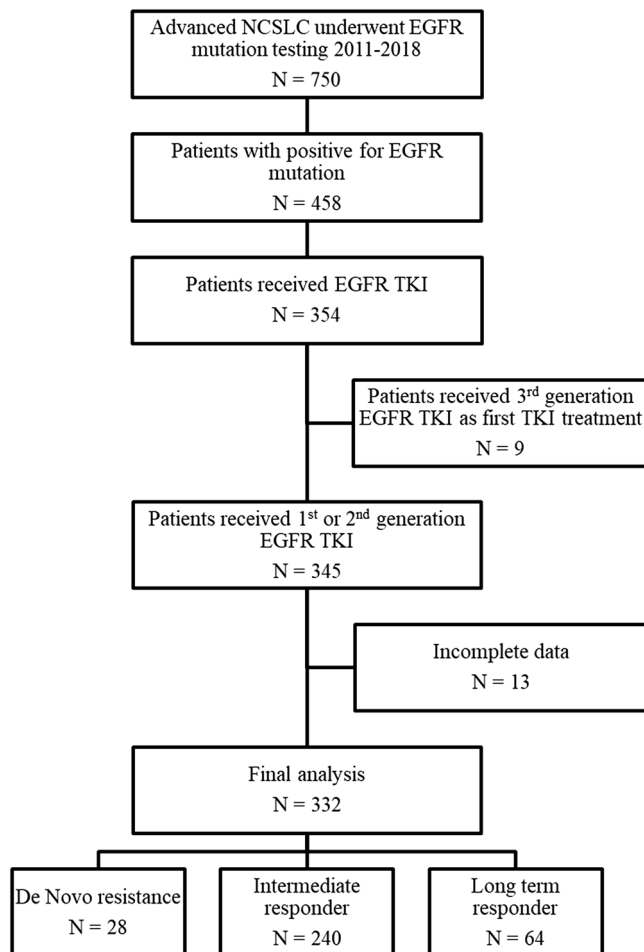
# Co-occurrence *CDK4/6* amplification serves as biomarkers of de novo EGFR TKI resistance in sensitizing *EGFR* mutation non-small cell lung cancer

Piyada Sitthideatphaiboon<sup>1</sup>, Chinachote Teerapakpinyo<sup>2</sup>, Krittiya Korphaisarn<sup>3</sup>, Nophol Leelayuwatanakul<sup>4</sup>, Nopporn Pornpatrananrak<sup>5</sup>, Naravat Pongvarin<sup>6</sup>, Poonchavist Chantranuwat<sup>7</sup>, Shanop Shuangshoti<sup>2,7</sup>, Chatchawit Apornthewan<sup>8</sup>, Wariya Chintanapakdee<sup>9</sup>, Virote Sriuranpong<sup>1</sup> & Chanida Vinayanuwattikun<sup>1✉</sup>

Despite the development of predictive biomarkers to shape treatment paradigms and outcomes, de novo EGFR TKI resistance advanced non-small cell lung cancer (NSCLC) remains an issue of concern. We explored clinical factors in 332 advanced NSCLC who received EGFR TKI and molecular characteristics through 65 whole exome sequencing of various EGFR TKI responses including; de novo (progression within 3 months), intermediate response (IRs) and long-term response (LTRs) (durability > 2 years). Uncommon *EGFR* mutation subtypes were significantly variable enriched in de novo resistance. The remaining sensitizing *EGFR* mutation subtypes (exon 19 del and L858R) accounted for 75% of de novo resistance. Genomic landscape analysis was conducted, focusing in 10 frequent oncogenic signaling pathways with functional contributions; cell cycle, Hippo, Myc, Notch, Nrf2, PI-3-Kinase/Akt, RTK-RAS, TGF- $\beta$ , p53 and  $\beta$ -catenin/Wnt signaling. Cell cycle pathway was the only significant alteration pathway among groups with the FDR  $p$ -value of  $6 \times 10^{-4}$ . We found only significant  $q$ -values of  $< 0.05$  in 7 gene alterations; *CDK6*, *CCNE1*, *CDK4*, *CCND3*, *MET*, *FGFR4* and *HRAS* which enrich in de novo resistance [range 36–73%] compared to IRs/LTRs [range 4–22%]. Amplification of *CDK4/6* was significant in de novo resistance, contrary to IRs and LTRs (91%, 27.9% and 0%, respectively). The presence of co-occurrence *CDK4/6* amplification correlated with poor disease outcome with HR of progression-free survival of 3.63 [95% CI 1.80–7.31,  $p$ -value  $< 0.001$ ]. The presence of *CDK4/6* amplification in pretreatment specimen serves as a predictive biomarker for de novo resistance in sensitizing *EGFR* mutation.

The discovery of molecularly targeted therapies has dramatically changed the paradigm of treatment in advanced non-small cell lung cancer (NSCLC). Activating epidermal growth factor receptor (*EGFR*) mutations are key

<sup>1</sup>Division of Medical Oncology, Department of Medicine, Faculty of Medicine, Chulalongkorn University and The King Chulalongkorn Memorial Hospital, Bangkok 10330, Thailand. <sup>2</sup>Chula GenePRO Center, Research Affairs, Chulalongkorn University and The King Chulalongkorn Memorial Hospital, Bangkok 10330, Thailand. <sup>3</sup>Division of Medical Oncology, Department of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Siriraj, Bangkok Noi, Bangkok 10700, Thailand. <sup>4</sup>Division of Pulmonary and Critical Care Medicine, Department of Medicine, Faculty of Medicine, Chulalongkorn University and The King Chulalongkorn Memorial Hospital, Bangkok, Thailand. <sup>5</sup>Department of Surgery, Faculty of Medicine, Chulalongkorn University and The King Chulalongkorn Memorial Hospital, Bangkok 10330, Thailand. <sup>6</sup>Department of Clinical Pathology, Faculty of Medicine Siriraj Hospital, Mahidol University, Siriraj, Bangkok Noi, Bangkok 10700, Thailand. <sup>7</sup>Department of Pathology, Faculty of Medicine, Chulalongkorn University and The King Chulalongkorn Memorial Hospital, Bangkok 10330, Thailand. <sup>8</sup>Department of Mathematics and Computer Science & Omics Science and Bioinformatics Center, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand. <sup>9</sup>Department of Radiology, Faculty of Medicine, Chulalongkorn University and The King Chulalongkorn Memorial Hospital, Bangkok, Thailand. ✉email: Chanida.Vi@chula.ac.th



**Figure 1.** Consort diagram of 332 EGFR mutation-positive recurrence or advanced NSCLC patients in this study.

drivers of NSCLC and are more common in East Asian than non-Asian populations (40–60% vs. 10–15%, respectively)<sup>1–3</sup>. Several randomized studies support the use of EGFR tyrosine kinase inhibitors (TKIs) as the standard first-line treatment for patients with activating *EGFR* mutations<sup>4–11</sup>. However, there is heterogeneity in treatment responses of EGFR TKIs with a progression-free survival (PFS) ranging from a few months to several years and resistance disease inevitably emerges. Approximately 10–20% of patients that harbor activating *EGFR* mutations do not exhibit objective responses to EGFR TKIs. Previous studies have identified molecular mechanisms that are associated with diverse responses to TKIs, including *EGFR* mutation subtype<sup>12,13</sup>, primary existence of the *EGFR* T790M mutation or coexistence of *EGFR* mutations with other genetic alterations<sup>14,15</sup>, for example, *KRAS* mutation, *MET* amplification, *PIK3CA* mutation, inactivation of *TP53*, or *BIM* polymorphism. Taken together, comparing clinical characteristics and comprehensive genomic landscapes via whole exome sequencing (WES) in diverse EGFR TKI responders; de novo resistance, intermediate and long-term responses, *EGFR* mutation NSCLC will enable the identification of potential mechanisms that confer resistance to EGFR TKI treatment.

## Results

**Patient demographics.** Of the 458 patients with NSCLC whose tumors harbored *EGFR*-activating mutation, 332 patients received EGFR TKIs and complete follow-up data were included in the final analysis (Fig. 1). Patient demographics are summarized in Table 1. The median age was 64 years (interquartile range [IQR] 54.3 to 72 years). Sixty-four percent of patients were women. Most of the patients were never smokers (80%) and had 0–1 score of ECOG performance status (87%). A majority of the patients were adenocarcinoma (95%), metastatic disease at presentation (76%) and 1–2 metastatic sites (75%). Baseline brain metastases were present in 22% of the overall population. Regarding *EGFR* mutation subtype, 169 patients (51%) harbored exon 19 deletion, 136 patients (41%) harbored L858R, and 27 patients (8%) had other mutations including G719X in exon 18 (N=6), exon 20 insertion (N=2), S768I in exon 20 (N=1), and L861G or Q in exon 21 (N=6). Twelve patients had any or two coexisting *EGFR* mutations (complex mutation) including one patient with a L858R with a coexisting de novo *EGFR* T790M mutation, respectively. A total of 218 patients (66%) received EGFR TKIs as first-line treatment, and the remaining were treated as the subsequent line treatment. The majority of 1st generation EGFR TKIs were administered up to 95%, composed of gefitinib 59% and erlotinib 36%. The objective response rate

Characteristics	All (N = 332)	De novo resistance (N = 28)	Intermediate responders (N = 240)	Long-term responders (N = 64)	P values
<b>Age at diagnosis, n (%)</b>					0.44
Median (IQR)	64 (54.3–72)	60.5 (50–68)	64 (55–72)	63 (55–72)	
< 60 years	136 (41%)	14 (50%)	99 (41.3%)	23 (35.9%)	
≥ 60 years	196 (59%)	14 (50%)	141 (58.7%)	41 (64.1%)	
<b>Gender, n (%)</b>					0.56
Male	118 (35.5%)	12 (42.9%)	86 (35.8%)	20 (31.3%)	
Female	214 (64.5%)	16 (57.1%)	154 (64.2%)	44 (68.8%)	
<b>ECOG PS, n (%)</b>					0.07
0–1	272 (86.6%)	21 (75%)	197 (86.4%)	54 (93.1%)	
≥ 2	42 (13.4%)	7 (25%)	31 (13.6%)	4 (6.9%)	
Missing	18		12	6	
<b>Smoking status, n (%)</b>					0.10
Never	229 (80.4%)	16 (64%)	168 (82%)	45 (81.8%)	
Current/former	56 (19.6%)	9 (36%)	37 (18%)	10 (18.2%)	
Missing	47	3	35	9	
<b>Histology, n (%)</b>					0.48
Adenocarcinoma	310 (95.4%)	28 (100%)	225 (95%)	57 (95%)	
Non adenocarcinoma	15 (4.6%)	0	12 (5%)	3 (5%)	
Missing	7		3	4	
<b>Stage at diagnosis, n (%)</b>					<b>0.003*</b>
Recurrent	79 (23.8%)	1 (3.6%)	55 (22.9%)	23 (35.9%)	
Metastatic	253 (76.2%)	27 (96.4%)	185 (77.1%)	41 (64.1%)	
Curative surgery, n (%)	70 (21.1%)	1 (3.6%)	48 (20%)	21 (32.8%)	<b>0.01*</b>
<b>Number of metastatic sites, n (%)</b>					<b>0.002*</b>
1–2 sites	248 (74.7%)	20 (71.4%)	169 (70.4%)	59 (92.2%)	
≥ 3 sites	84 (25.3%)	8 (28.6%)	71 (29.6%)	5 (7.8%)	
Brain metastasis, n (%)	73 (22%)	9 (32.1%)	54 (22.5%)	10 (15.6%)	0.20
Liver metastasis, n (%)	37 (11.1%)	2 (7.1%)	29 (12.1%)	6 (9.4%)	0.65
<b>EGFR mutation subtypes, n (%)</b>					<b>&lt;0.001*</b>
Del19	169 (50.9%)	10 (35.7%)	126 (52.5%)	33 (51.6%)	
L858R	136 (41%)	10 (35.7%)	98 (40.8%)	28 (43.8%)	
G719X	6 (1.8%)	3 (10.7%)	3 (1.3%)	0	
L861G or Q	6 (1.8%)	3 (10.7%)	3 (1.3%)	0	
S768I	1 (0.3%)	0	1 (0.4%)	0	
Exon 20 insertion	2 (0.6%)	1 (3.6%)	1 (0.4%)	0	
Any or combined mutations	12 (3.6%)	1 (3.6%)	8 (3.3%)	3 (4.7%)	
<b>EGFR mutation subtypes, n (%)</b>					<b>&lt;0.001*</b>
Del19	169 (50.9%)	10 (35.7%)	126 (52.5%)	33 (51.6%)	
L858R	136 (41%)	10 (35.7%)	98 (40.8%)	28 (43.8%)	
Uncommon <sup>a</sup>	15 (4.5%)	7 (25%)	8 (3.3%)	0	
Any or combined mutations	12 (3.6%)	1 (3.6%)	8 (3.3%)	3 (4.7%)	
<b>Line of EGFR TKIs, n (%)</b>					0.21
First line	218 (65.7%)	17 (60.7%)	153 (63.8%)	48 (75%)	
Later line	114 (34.3%)	11 (39.3%)	87 (36.3%)	16 (25%)	
<b>Type of EGFR TKIs, n (%)</b>					<b>0.01*</b>
Gefitinib	197 (59.3%)	17 (60.7%)	145 (60.4%)	35 (54.7%)	
Erlotinib	121 (36.4%)	8 (28.6%)	84 (35%)	29 (45.3%)	
2nd generation <sup>b</sup>	14 (4.3%)	3 (10.7%)	11 (4.6%)	0	
Duration of TKI treatment, median (IQR, months)	12.4 (7.1–20.5)	2.3 (1.6–3)	11.9 (7.6–15.9)	32.4 (27.8–37.3)	<b>&lt;0.001*</b>
ORR, n (%)	201 (60.5%)	1 (3.6%)	159 (66.2%)	41 (64.1%)	<b>&lt;0.001*</b>

**Table 1.** Patient demographics in the overall population, 332 advanced or recurrent NSCLC who received 1st or 2nd generation EGFR TKI. <sup>a</sup>Uncommon EGFR mutations, including G719X in exon 18 (n = 6), exon 20 insertion (n = 2), S768I in exon 20 (n = 1) and L861G or Q in exon 21 (n = 6). <sup>b</sup>Afatinib and dacomitinib were used in 13 and 1 patients, respectively. *Del19* exon 19 deletion, *ECOG PS* Eastern Cooperative Oncology Group performance status, *EGFR* epidermal growth factor receptor, *IQR* interquartile range, *ORR* overall response rate.

(ORR) of EGFR TKIs in our cohort was 60.5%. The median duration of EGFR TKIs treatment was 12.4 months (IQR 7.1 to 20.5 months), the median duration of response was 10.0 months (IQR, 5.6–16.3 months), the median duration of stable disease (SD) was 9.1 months (IQR, 4.7–16.3 months), and the median PFS for any lines of EGFR-TKI was 12.2 months (95% confidence interval [CI] 11.0–13.3 months). At data cut-off on June 30, 2020, the median follow-up duration was 51.5 months (95% CI 45.9–57.0). A total of 217 patients (65%) ceased. The median OS of the overall study cohort was 32.9 months (95% CI 27.9–37.8 months).

**Clinical factors associated with the response of EGFR TKIs.** Patients were categorized into 3 groups based on EGFR TKI responsiveness (1) de novo resistance, EGFR TKI resistance who were defined as the best response was progressive disease (PD) or SD less than 3 months while receiving EGFR TKIs. This group represented intrinsic resistance to EGFR TKIs<sup>16</sup>; (2) intermediate responder, developed acquired resistance to EGFR TKIs according to the proposed criteria by Jackman<sup>17</sup> and (3) long-term responder, durable disease control with EGFR TKIs more than 2 years<sup>18</sup>. Patient characteristics of the three groups were listed in Table 1.

The de novo resistance groups were significantly associated with metastatic disease at presentation (96.4%;  $p$ -value 0.003) and uncommon *EGFR* mutation subtype (25%;  $p$ -value 0.001). Presence of metastatic disease at diagnosis was found in 77% and 64% of IRs and LTRs, respectively. Only 3.3% of tumors harboring uncommon *EGFR* mutations were found in IRs and absent in LTRs. There was no significant difference in age, gender, PS, smoking status, histology, baseline liver or brain metastasis, and the treatment lines of TKI between the three groups. Moreover, there was no difference in EGFR TKIs response between the exon 19 deletion and L858R mutation.

Logistic regression was performed to evaluate the correlation of clinical variables and response to EGFR TKIs (Table 2, Table S1) which revealed uncommon *EGFR* mutation subtype as significant variables in de novo resistance vs. IRs and LTRs with odds ratios (OR) of 6.83 ([95% CI 2.36–19.80],  $p$ -value < 0.001) and 16.84 ([95% CI 1.66–171.45,  $p$ -value 0.02), respectively. Poor ECOG performance status and metastatic disease at presentation were independent factors in de novo resistance vs. LTRs with OR of 7.39 ([95% CI 1.16–47.16,  $p$ -value = 0.04) and 61.45 ([95% CI 1.29–not estimated],  $p$ -value = 0.04), respectively (Table 2). These results were consistent with the results of multivariate Cox's proportional hazards analysis which revealed that those factors were correlated with PFS of EGFR TKIs (Table S2). Kaplan–Meier analysis, according to each clinical variable, correlated with PFS/OS and results were shown in Fig. S1 and Fig. S2. Survival analysis and subsequent treatment were described in the supplementary information (Table S3–S6).

**Comparative “cohort-normal” vs. “match-normal” WES analysis workflow in exploratory cohort 65 tumor-normal resectable lung cancer.** To define the concordance variant calling between “cohort-normal” and “match-normal”, we conducted WES analysis in 65 patients with resectable stage adenocarcinoma of the lung who underwent surgery as a curative intent. A “cohort-normal” pipeline was conducted using in-house normal reference obtained from either leucocyte or normal lung. In general, mutation profiling for 21 driver genes and CNAs (arm-level and focal-level) were consistent with the lung adenocarcinoma East Asian cohort<sup>19</sup> (Fig. S5). The most frequent driver mutations were *EGFR* (60%), *TP53* (28%) and *RMB10* (11%), consistent with the East Asian cohort (47%, 36% and 8%, respectively), while the *KRAS* mutation was found 4% lower than the East Asian cohort (11%). Median TMB (including synonymous and non-synonymous mutations) was low at 1.84 Mb<sup>-1</sup> (range: 0.24–25.14 Mb<sup>-1</sup>) which is a dominant characteristic of the majority of never smoker, adenocarcinoma lung cancer (73%) in our study. Many focal CNAs were found around driver gene amplification in *EGFR*, *MYC*, *MDM2*, *KRAS* and *CCNE1* as well as deletion in *ARID1A* and *APC* (Fig. S6A and S6B). Somatic prediction in “cohort-normal” workflow was conducted using PureCN based on altered allelic fractions of germline and somatic variants which previously showed median accuracy of somatic variants of 97.2% in TCGA-LUAD<sup>20</sup>. There was a high correlation ( $R = 0.99$ ,  $p$ -value <  $2.2 \times 10^{-16}$ ) of all non-synonymous mutations between “cohort-normal” and “match-normal” workflow (Fig. S7B). There were 3,445 non-synonymous variants in the “cohort-normal” workflow and 4,717 non-synonymous variants in the “match-normal” workflow. Eighty-four percent of all non-synonymous mutations in “cohort-normal” were concordant to 61.3% of “match-normal” workflow. The concordant rates in “cohort-normal” were 89% and 92.3% in 307 significant genes from LUAD 7 studies (additional information: Table S12) and 206 genes from 10 significant pathway analysis, respectively<sup>21</sup> (Fig. S7A). Cohort-normal workflow of the non-synonymous mutation variant with high-concordance rate to match-normal workflow in 206 genes was adopted in WES analysis of 65 *EGFR* mutation-positive recurrence or advanced NSCLC. The retained variants according to the filtered algorithm is shown in Fig. S8. Demographic characteristics of 65 resectable stage adenocarcinoma of lung were shown in additional information: Table S11.

**Elucidation of molecular analysis correlation with the response of EGFR TKI in “cohort-normal”.** We selected participants for genomic study based on retrospective aforementioned-response classification. Demographic characteristics, response treatment of 65 advanced stage NSCLC, received EGFR TKIs who had adequate tissue for WES were shown in Table 3, Fig S3. We analyzed exome sequencing with target sequences of approximately 90 Mb. The average depth of coverage within targets was 65 $\times$  (range 60–94 $\times$ ) with 95% of targeted bases were covered by at least 10 reads. Based on the “cohort-normal” algorithm, 14,508 non-synonymous variants were retained from 65 WES recurrence or advanced *EGFR* mutation-positive NSCLC (additional information: Table S7). Median non-synonymous mutation was 2.3 Mb<sup>-1</sup> (range 1.5–6.0 Mb<sup>-1</sup>). The median frequency of non-synonymous mutation in de novo resistance was 1.15 Mb<sup>-1</sup> (range 0.65–3.33 Mb<sup>-1</sup>) lower than IRs and LTRs which was 2.82 Mb<sup>-1</sup> (range 1.07–6.02 Mb<sup>-1</sup>,  $p$ -value < 0.001) and 1.77 Mb<sup>-1</sup> (range 1.18–2.98,  $p$ -value 0.01). However, this might be an effect of lower average read depth in de novo resistance than

De novo resistance versus IRs variables <sup>a</sup>	Univariate		Multivariate	
	OR (95% CI)	P value	OR (95% CI)	P value
Age (< 60/≥ 60)	1.42 (0.65–3.12)	0.38		
Sex (male/female)	1.34 (0.61–2.97)	0.47		
ECOG PS (≥ 2/0–1)	2.12 (0.83–5.40)	0.12		
Smoking (current-former/never)	2.55 (1.05–6.23)	<b>0.04*</b>	2.14 (0.83–5.54)	0.12
Histology (non-ADC/ADC)	Not estimated	1		
Stage at diagnosis (M1/M0)	8.03 (1.07–60.42)	<b>0.04*</b>	5.01 (0.65–38.91)	0.12
Curative surgery (no/yes)	3.25 (0.75–14.17)	0.12		
Number of metastatic sites (≥ 3/1–2)	0.95 (0.40–2.26)	0.91		
Brain metastasis (yes/no)	1.63 (0.70–3.81)	0.26		
Liver metastasis (yes/no)	0.56 (0.13–2.48)	0.45		
EGFR subtype (others/common)	5.60 (2.14–14.69)	<b>&lt;0.001*</b>	6.83 (2.36–19.80)	<b>&lt;0.001*</b>
Common EGFR subtype (Del19/L858R)	0.78 (0.31–1.94)	0.59		
Line of TKI (first/late)	0.88 (0.39–1.96)	0.75		
Generation of TKI (first/second)	0.40 (0.11–1.53)	0.18		
First generation of TKI (gefitinib/erlotinib)	1.23 (0.51–2.98)	0.64		
De novo resistance versus LTRs variables <sup>a</sup>	Univariate		Multivariate	
	OR (95% CI)	P value	OR (95% CI)	P value
Age (< 60/≥ 60)	1.78 (0.73–4.38)	0.21		
Sex (male/female)	1.65 (0.66–4.13)	0.28		
ECOG PS (≥ 2/0–1)	4.50 (1.19–16.98)	<b>0.03*</b>	7.39 (1.16–47.16)	<b>0.04*</b>
Smoking (current-former/never)	2.53 (0.87–7.35)	0.09	2.38 (0.55–10.33)	0.25
Histology (non-ADC/ADC)	Not estimated	1		
Stage at diagnosis (M1/M0)	15.15 (1.93–118.86)	<b>0.01*</b>	61.45 (1.29–NE)	<b>0.04*</b>
Curative surgery (no/yes)	6.35 (1.38–29.32)	<b>0.02*</b>	0.30 (0.02–5.55)	0.42
Number of metastatic sites (≥ 3/1–2)	4.72 (1.38–16.10)	<b>0.01*</b>	4.05 (0.54–30.53)	0.18
Brain metastasis (yes/no)	2.56 (0.90–7.25)	0.08	1.92 (0.42–8.67)	0.40
Liver metastasis (yes/no)	0.74 (0.14–3.93)	0.73		
EGFR subtype (others/common)	8.13 (1.97–33.64)	<b>0.004*</b>	16.84 (1.66–171.45)	<b>0.02*</b>
Common EGFR subtype (Del19/L858R)	0.85 (0.31–2.33)	0.75		
Line of TKI (first/late)	0.52 (0.20–1.33)	0.17		
Generation of TKI (first/second)	Not estimated	1		
First generation of TKI (gefitinib/erlotinib)	1.76 (0.67–4.66)	0.26		

**Table 2.** Univariate and multivariate analyses of clinical variables and response to EGFR TKIs. <sup>a</sup>Category after the slash (/) was set as reference category. ADC adenocarcinoma, Del19 exon 19 deletion, ECOG PS Eastern Cooperative Oncology Group performance status, EGFR epidermal growth factor receptor, IRs intermediate responders, LTRs long-term responders, M0 recurrent disease, M1 metastatic disease, TKI tyrosine kinase inhibitor.

IRs and LTRs, 53× (range 40–63×), 68× (range 60–94×,  $p$ -value < 0.001) and 64.3× (range 40–80×,  $p$ -value 0.03) (Fig. S9A). No statistical difference was present in tumor ploidy nor tumor purity among 3 groups. Median tumor ploidies were 3.6 (range 1.9–4.6) in de novo resistance, 2.4 (range 1.0–4.6,  $p$ -value 0.058) in IRs and 2.6 (range 1.6–3.8,  $p$ -value 0.055) in LTRs (Fig. S9B). Median tumor purities were 0.38 (range 0.23–0.59) in de novo resistance, 0.33 (range 0.18–0.66,  $p$ -value = 0.33) in IRs and 0.37 (range 0.26–0.61,  $p$ -value = 0.94) in LTRs (Fig. S9C). Significant CNAs of advanced 65 WES recurrence or advanced NSCLC and gene-level segment integer copies number estimation are shown in additional information: Table S8.

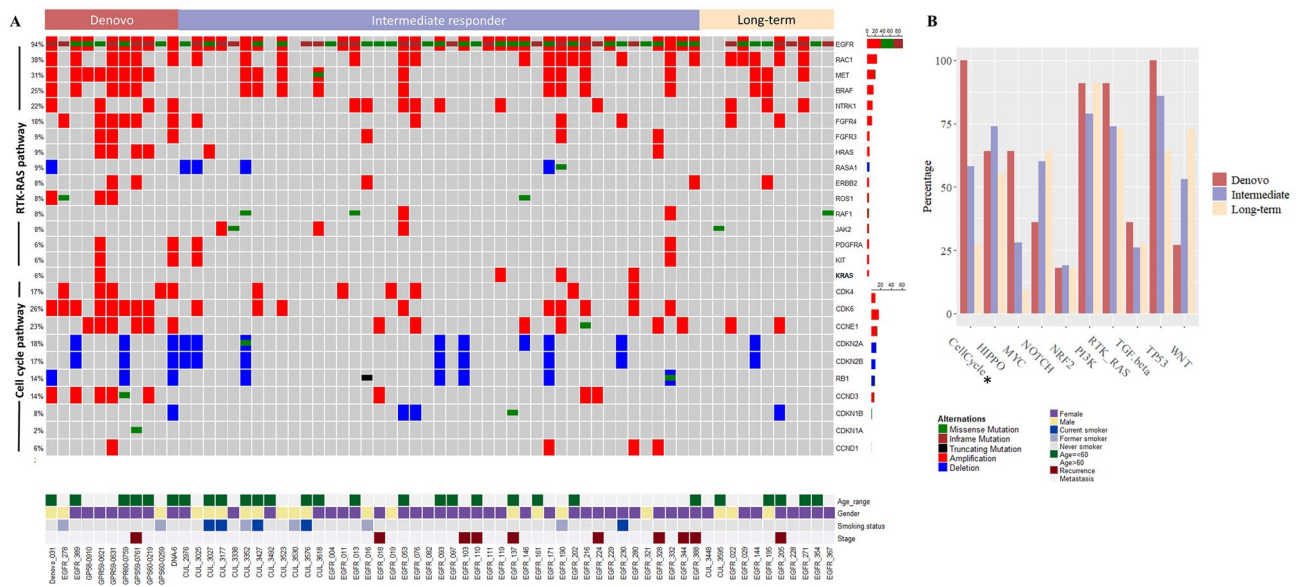
To define underlined molecular characteristics according to the response of treatment, we explored 10 oncogenic signaling pathways which frequent genetic alterations and might be candidates of functional contributions including; cell cycle, Hippo, Myc, Notch, Nrf2, PI-3-Kinase/Akt, RTK-RAS, TGF- $\beta$ , p53 and  $\beta$ -catenin/Wnt signaling<sup>21</sup>. A list of genes in each pathway and pattern of alterations are shown in additional information: Table S9, Fig. S10A–S10J. A tumor with one or more gene alterations, either mutation or CNAs in the pathway, was considered pathway alterations. The frequency of 10 pathway alterations was in range the of 18–84% (18.4% Nrf2, 27.6% TGF- $\beta$ , 30% Myc, 52.3%  $\beta$ -catenin/Wnt, 56% Notch, 60% cell cycle, 76.9% RTK-RAS, 69% Hippo, 83% PI-3-Kinase/Akt, and 84% p53 pathway). In general, the prevalence of individual genetic alterations in our study were; TP53 mutation 60%, MDM2 copy number gain (CNG) 15%, MET CNG 30%, ERBB2 CNG 7.6%, PIK3CA alteration 16%, CDK4 CNG 16.9%, CDK6 CNG 26.1%, RB1 alteration 13.8%, CCNE1 CNG 21%. The majority of frequency alterations, except CDK6 and CCNE1, were similar to the EGFR mutation-positive NSCLC literature<sup>22</sup>.

Cell cycle, RTK-RAS and PI-3-Kinase/Akt were the significant alteration pathways among treatment groups with the  $p$ -value of  $6 \times 10^{-5}$ , 0.02 and 0.02, respectively. The  $P$ -value for significant pathways were adjusted by

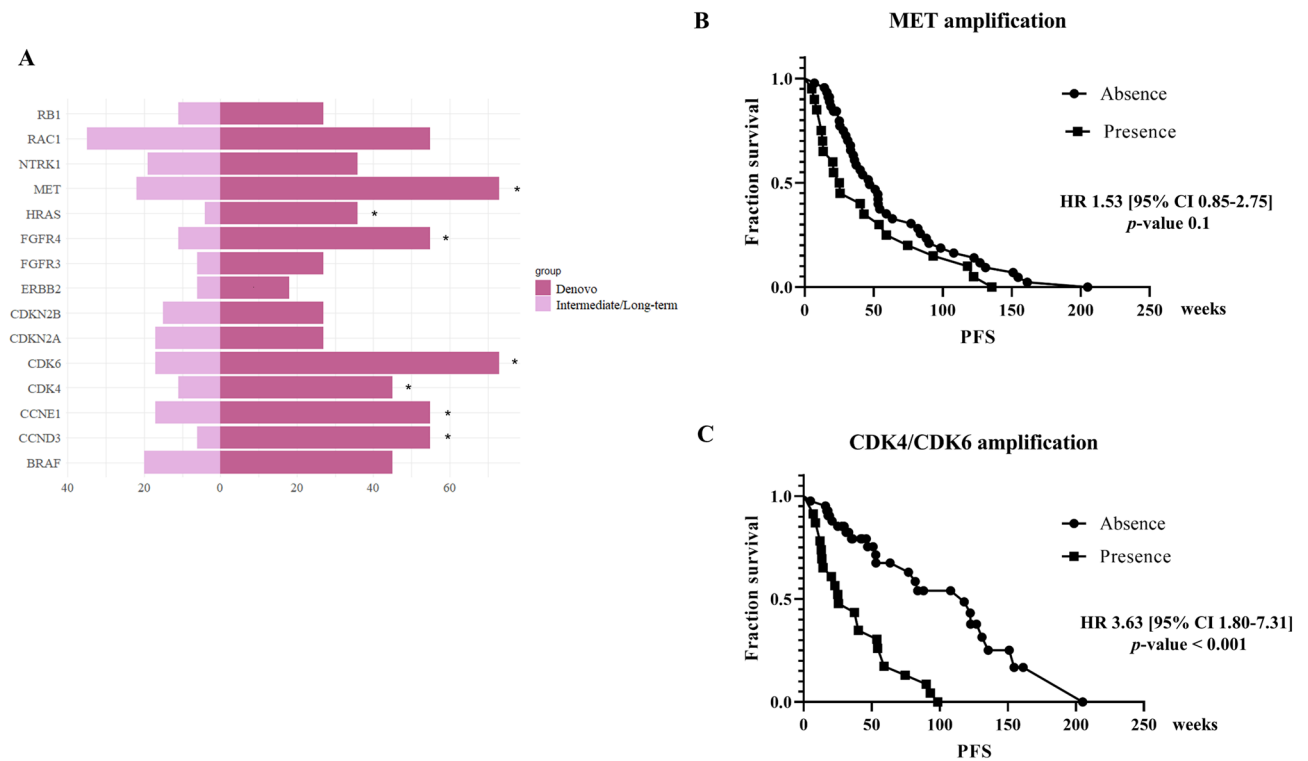
Characteristics	All (N = 65)	De novo resistance (N = 11)	Intermediate responders (N = 44)	Long-term responders (N = 10)
<b>Age at diagnosis, n (%)</b>				
< 60 years	28 (43%)	6 (54.6%)	18 (40.9%)	4 (40%)
≥ 60 years	37 (57%)	5 (45.4%)	26 (59.1%)	6 (60%)
<b>Gender, n (%)</b>				
Male	20 (30.7%)	3 (27.2%)	15 (34%)	2 (20%)
Female	45 (69.3%)	8 (72.8%)	29 (66%)	8 (80%)
<b>Smoking status, n (%)</b>				
Never	54 (83%)	9 (81.9%)	35 (79.5%)	10 (100%)
Current/former	11 (17%)	2 (18.1%)	9 (20.5%)	0
<b>Histology, n (%)</b>				
Adenocarcinoma	61 (93.8%)	11 (100%)	40 (90.9%)	10 (100%)
NSCLC NOS	4 (6.2%)	0	4 (9.1%)	0
<b>Stage at diagnosis, n (%)</b>				
Recurrent	10 (15.3%)	1 (9.1%)	8 (18.2%)	1 (10%)
Metastatic	55 (84.7%)	10 (90.9%)	36 (81.8%)	9 (90%)
<b>EGFR mutation subtypes, n (%)</b>				
Del19	36 (55.3%)	7 (63.6%)	22 (50%)	7 (70%)
L858R	28 (43%)	3 (27.2%)	22 (50%)	3 (30%)
G719X	1 (1.7%)	1 (9.2%)	0	0
<b>Line of EGFR TKIs, n (%)</b>				
First line	49 (75.3%)	6 (54.5%)	34 (77.2%)	9 (81.8%)
Later line	16 (24.7%)	5 (45.5%)	9 (22.8%)	2 (18.2%)
<b>EGFR mutation subtypes, n (%)</b>				
Del19	36 (55.3%)	7 (63.6%)	22 (50%)	7 (70%)
L858R	28 (43%)	3 (27.2%)	22 (50%)	3 (30%)
Uncommon <sup>a</sup>	1 (1.7%)	1 (9.2%)	0	0
<b>Type of EGFR TKIs, n (%)</b>				
Gefitinib	38 (58.4%)	8 (72.7%)	26 (59%)	4 (40%)
Erlotinib	22 (33.8%)	2 (18.1%)	14 (31.8%)	6 (60%)
Afatinib	5 (7.8%)	1 (9.2%)	4 (9.2%)	0
Duration of TKI treatment, median (IQR, months)	9.3 (4.8–19.1)	2.8 (1.8–3.2)	9.6 (5.9–13.8)	31.1 (28.8–35.9)

**Table 3.** Demographic characteristic of 65 advanced or recurrence NSCLC participants who performed WES. *Del19* exon 19 deletion, *EGFR* epidermal growth factor receptor, *IQR* interquartile range.

Benjamini–Hochberg Method which revealed significantly less than 0.05 in only cell cycle pathways ( $q$ -value  $6 \times 10^4$ ) (Fig. 2B). Individual genetic alterations per pathway were shown in Fig. 2A, Fig. S10A–S10J and additional information: Table S10. Either the amplification of *CDK4* or *CDK6* was found significant in the de novo resistance group, contrary to IRs and LTRs (91%, 27.9% and 0%, respectively). No correlation between significant alteration in cell cycle, RTK-RAS pathway and clinical phenotypes such as age, sex, smoking status, stage of disease was present. Among those significant pathways, we found a significant  $q$ -value  $< 0.05$  in only 7 gene alterations; *CDK6*, *CCNE1*, *CDK4*, *CCND3*, *MET*, *FGFR4* and *HRAS* (including amplification and variant mutation) which enrich in de novo resistance [range 36–73%] compared to IRs/LTRs [range 4–22%] (Fig. 3A). Co-occurrence of *MET* amplification favors shorter EGFR TKI disease control than the absence of *MET* amplification with a median PFS of 25 weeks vs. 47 weeks (HR 1.53 [95% CI 0.85–2.75,  $p$ -value 0.1] (Fig. 3B). The discriminative effect was significant in the presence of either *CDK4* or *CDK6* amplification. Median PFS of EGFR TKI was 25 weeks compared to 118 weeks in the absence of those molecular markers with the HR of progression-free survival 3.63 [95% CI 1.80–7.31,  $p$ -value  $< 0.001$ ] (Fig. 3C). For individual analysis, the presence of *CDK6* amplification was shown to shorten PFS of EGFR TKI with the HR of PFS 2.22 [95% CI 1.24–4.0,  $p$ -value 0.007]. The presence of *CDK4* amplification (16%), with less prevalence than *CDK6* amplification (26%), was also shown a trend of shortened PFS of EGFR TKI with the HR of PFS 1.79 [95% CI 0.92–3.48,  $p$ -value 0.08] (Fig S4). Furthermore, there was no *CDK4/CDK6* amplification in the long-term EGFR TKI responder. Median *CDK4* and *CDK6* integer copies number estimation by PureCN were 7 [range 2–244] and 5 [range 2–9]. The validation of *CDK4/6* amplification was conducted hybrid capture-based NGS OrigiMed Gene variation testing kit (ONCOSNAP pro)<sup>23</sup> in 5 adequate specimens. The results were shown in additional information table S13. Consistent *CDK4* amplification in all but not for *CDK6* amplification was found. As OrigiMed Gene amplification threshold for amplification was over 6 copies. Five specimens which had *CDK6* amplification in range 2.9–6 using OrigiMed had been excluded by the algorithm. We analyzed the discriminative effect varying the WES threshold of amplification by using



**Figure 2.** Individual details of patient factors with genomic alterations including SNVs and CNAs in RTK-RAS and cell cycle pathway (A). The cell cycle pathway was the only pathway that showed statistically significant ( $q$ -value  $6 \times 10^{-4}$ ) results among 10 oncogenic signaling pathways (B). The frequency of cell cycle genomic alterations was lowest, 27% in LTRs, 58% in IRs and enriched (100%) in de novo resistance.



**Figure 3.** Comparison of 15 gene alteration frequencies among de novo resistance ( $n = 11$ ) and IRs/LTRs ( $n = 54$ ). We found a significant  $q$ -value  $< 0.05$  in only 7 gene alterations; *CDK6*, *CCNE1*, *CDK4*, *CCND3*, *MET*, *FGFR4* and *HRAS* (including amplification and variant mutation) which enrich in de novo resistance [range 36–73%] compared to IRs/LTRs [range 4–22%] (A). Co-occurrence of MET amplification favors shorter EGFR TKI disease control than the absence of MET amplification with a median PFS 25 weeks vs. 47 weeks (HR 1.53 [95% CI 0.85–2.75,  $p$ -value 0.1] (B). While the presence of either *CDK4* or *CDK6* amplification significantly correlated with de novo resistance with the HR of PFS 0.63 [95% CI 1.80–7.31,  $p$ -value  $< 0.001$ ] (C).

a calculated integer copy number. The HR of PFS were 1.92 [95% CI 1.09–3.41,  $p$ -value 0.02] and 2.17 [95% CI 1.13–4.15,  $p$ -value 0.01] for an amplification threshold of 4 and 6 respectively.

**Multivariate Cox regression analysis of PFS and OS including clinico-genomic characteristics.** The median PFS and OS of 65 patients who had available FFPE to perform WES were 10 months [range 1.2–51.2] and 22.1 months [range 3.2–103.2] respectively. We conducted multivariate Cox regression analysis including sex, age, ECOG PS, smoking status, stage and molecular alterations including *CDK4/6* amplification, *MET* amplification, *TP53* mutation and *EGFR* mutation subtype. We found that uncommon *EGFR* mutation (L861Q) HR 14.61 [95% CI 1.45–146.9,  $p$ -value = 0.023], current/former smoking status HR 5.48 [95% CI 1.87–16.1,  $p$ -value = 0.002] and presence of *CDK4/6* amplification HR 2.04 [95% CI 1.06–3.9,  $p$ -value = 0.03] were significantly associated with shorten PFS (Fig. 4A). Regarding OS, we added total regimen of treatment into aforementioned multivariate Cox regression analysis. We found that uncommon *EGFR* mutation (L861Q) HR 14.79 [95% CI 1.43–152.7,  $p$ -value = 0.024], current/former smoking status HR 7.03 [95% CI 2.1–23.0,  $p$ -value = 0.001] were significantly associated with OS, consistent with PFS. Furthermore, presence of *MET* amplification HR 2.13 [95% CI 1.04–4.5,  $p$ -value = 0.03] and presence of *TP53* mutation HR 2.06 [95% CI 0.93–4.6,  $p$ -value = 0.07] were also associated with shorten OS (Fig. 4B).

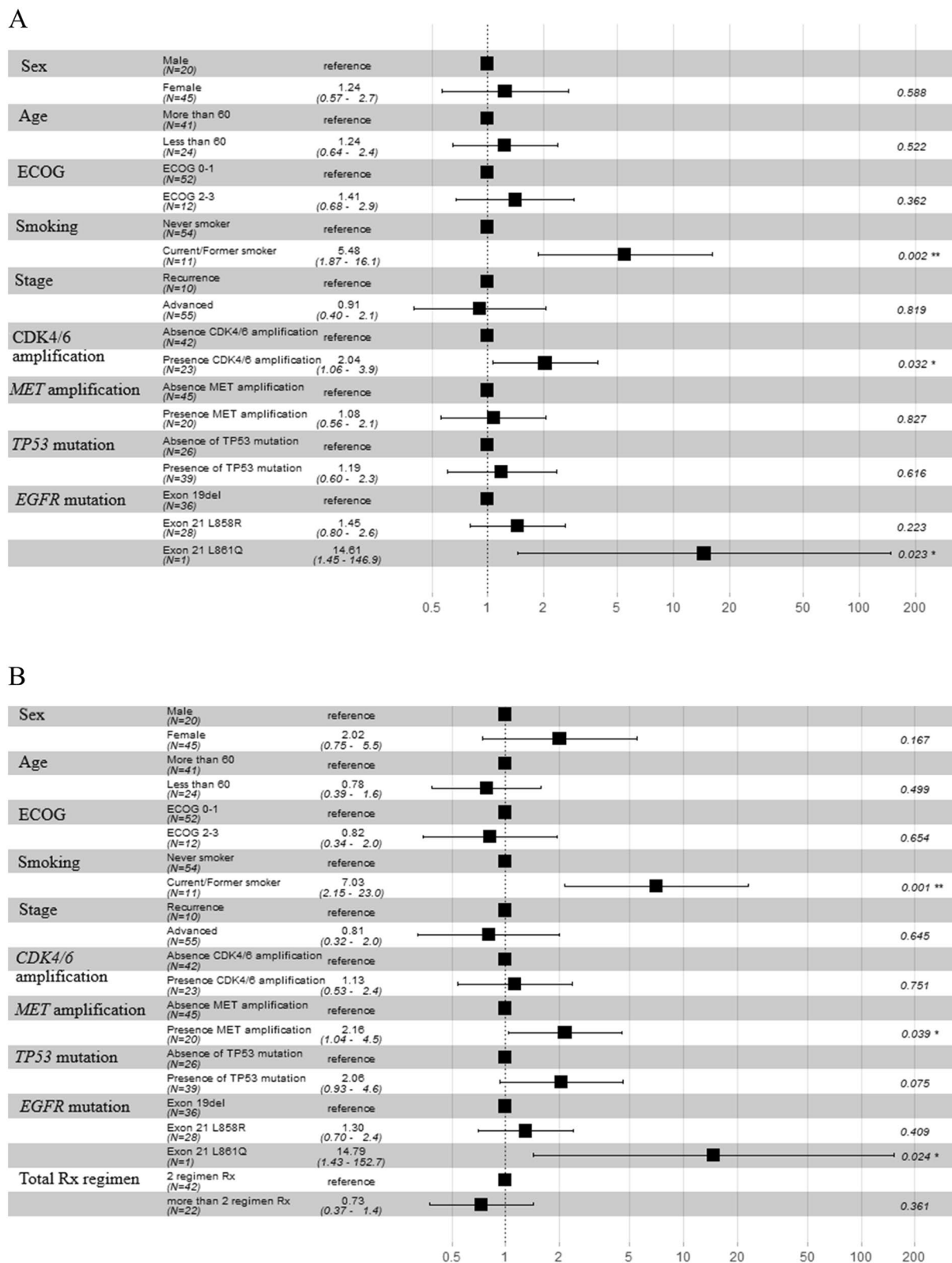
## Discussion

We analyzed the clinicopathological and molecular features of a subset of patients that is intrinsically resistant to EGFR TKI treatment, although this subset represented only 8% of our study population. We found that uncommon *EGFR* mutation was an independent factor associated with de novo resistance compared to both IRs and LTRs. It is well known that uncommon *EGFR* mutations are a heterogeneous group with variable responses to EGFR TKIs, contrary to LTRs which represented 19% of our study. The median duration of EGFR TKIs treatment in this group was 32.4 months. LTRs showed a substantially lower number of metastatic sites ( $p$ -value 0.002) and almost exclusively oligo-progression. Patients with recurrent disease who displayed favorable responses to EGFR TKI treatment may partly be explained by their small disease burden and low tumor heterogeneity<sup>24,25</sup>. Additionally, a meta-analysis showed that never smokers had better PFS benefits than ever smokers in patients who harbored the activating *EGFR* mutation and received EGFR TKIs<sup>26</sup>. However, we found that this factor was not associated with the outcome of disease control.

Besides specific clinical factors, we found that diverse genomic landscapes underlined distinct EGFR TKI responses. Varying mechanisms of de novo resistance in sensitizing *EGFR* mutation were reported such as de novo co-occurrence of *PIK3CA*<sup>27</sup>, *PI3K/AKT/mTOR*<sup>28</sup>, *PTEN* loss<sup>29</sup>, *MET* alteration<sup>30–32</sup> and *TP53* mutation<sup>33</sup>. Here, we focused on 10 frequent oncogenic signaling pathways; cell cycle, Hippo, Myc, Notch, Nrf2, PI-3-Kinase/Akt, RTK-RAS, TGF- $\beta$ , p53 and  $\beta$ -catenin/Wnt signaling which were previously shown to be significant among various cancer types, involving tumorigenesis, cell proliferation, metastasis and angiogenesis<sup>34,35</sup>. Targeting signaling pathways has been a challenge in defining a novel cancer treatment. Among them, RTK-RAS and cell cycle pathways were the most frequent alterations in no-mutation selected adenocarcinoma of the lung with a frequency of 74% and 56%, respectively<sup>21</sup>. These frequencies of pathway alterations were consistent with our *EGFR* mutation-positive study; 77% in RTK-RAS and 60% in cell cycle pathway, respectively. We found that cell cycle pathway alteration was the only significant pathway alteration ( $q$ -value < 0.05) with essential frequency in de novo compared to IRs and LTRs (100% vs. 58% vs. 27%, respectively). *CCNE1*, *CDK4/6* and *CCND3* were major contributors of cell cycle pathway alteration with  $q$ -value < 0.05. Altered cell cycle expression has also been correlated with acquired EGFR TKI resistance<sup>36</sup>. Broadened exploration of gene alteration in our study confirmed prior cfDNA targeted sequencing of 68 genes that revealed significant cell cycle pathways and the presence of *CDK4/6* alterations, which were significantly associated with non-responder of osimertinib<sup>37</sup>.

Furthermore, *MET* alteration was also enriched in de novo resistance with  $q$ -value < 0.05. Despite varying techniques and definitions used to define *MET* amplification<sup>38</sup>, it is well-known as an important role in de novo and acquired EGFR TKI resistance through bypass activating ERBB/PI3K-Akt signaling pathway<sup>30,39–42</sup>. The presence of *MET* amplification was significantly associated with shortened OS in multivariate Cox regression analysis with the HR of 2.13 [95% CI 1.04–4.5,  $p$ -value = 0.03] (Fig. 4B). Missense *TP53* mutation, which has previously shown the predictive impact of EGFR TKI treatment in meta-analysis<sup>33</sup>, had a higher prevalence in de novo EGFR TKI than IRs and LTRs (81% vs. 60% vs. 36%, respectively). The prevalence of co-occurrence alteration of *TP53* and *CDK4/6* amplification was 24.6% (81% in de novo EGFR TKI, 16% in IRs and none in LTR). The presence of missense *TP53* mutation was shown as potential prognostic significance to OS but not for PFS in multivariate Cox regression analysis with the HR of OS 2.06 [95% CI 0.93–4.6,  $p$ -value = 0.07] (Fig. 4). Our results were consistent with previous publication<sup>43</sup>. The prevalence of co-alteration *RBI* alteration and *CDK4/6* amplification in *EGFR* mutation NSCLC patient was 7.6% (de novo 27%, IRs 4% and none in LTR). *RBI* alteration was correlated poor prognostic outcome in non-select genomic subgroup advanced NSCLC<sup>44</sup> but not in *EGFR*-selected population. Univariate analysis in *RBI* alteration was not shown significant in term of PFS and OS (HR of PFS 1.76 [95% CI 0.85–3.64,  $p$ -value 0.1], HR of OS 1.78 [95% CI 0.78–4.02,  $p$ -value 0.1]). Furthermore, non-canonical *CDK4/6* substrates such as transcription factor Forkhead Box 1 (*FOXM1*), certain glycolytic enzymes and nuclear factor of activated T cell (*NFAT*) family members, let activity of *CDK4/6* even lack of *RBI* function<sup>45–48</sup>. The presence of either *CDK4* or *CDK6* amplification in the pretreatment specimen served as a predictive biomarker for EGFR TKI resistance in sensitizing *EGFR* mutation. Correlation with calculated integer copy number, using *CDK4/6* amplification threshold either 2, 4, 6 has discriminate predictive significance to EGFR TKI. Combination EGFR TKI treatment plus anti-*CDK4/6* inhibitors are possible to overcome de novo EGFR TKI clonal resistance. Dual *CDK4/6* and *EGFR* blockage shown in vitro activity to prevent or delay resistance in *EGFR* mutant NSCLC<sup>49</sup>.





**Figure 4.** Multivariate analysis of PFS (A) and OS (B) in 65 advanced stage NSCLC who had adequate tissue for WES.

Although we focused on significant pathway alteration in whole exome sequencing, our study has some limitations. First, there was a lack of fusion alteration in our analysis. Fusion alteration has been reported as an uncommon mechanism of acquired EGFR TKI resistance<sup>50–53</sup>. Nevertheless, co-occurrence fusion in pretreatment EGFR mutation-positive was reported at a low frequency (0.9%)<sup>52</sup>. Second, we used cohort-normal workflow, which is required in silico prediction by using allele-specific copy number to calculate the posterior probability to define the variant as somatic variant status (See “Methods”). Using an exploratory cohort of 65 tumor-normal pair resectable adenocarcinoma of lung, revealed that 84% of all non-synonymous mutations in “cohort-normal” concordance to 61.3% of “match-normal” workflow. Subclonal mutations which had low allelic fractions or low purity might be the reason for lower precision accuracy. Nevertheless, we selected high concordance, 92.3% in a limited 206 genes from 10 significant pathway analyses, which cover all significant co-occurrence alterations. Third, our average depth coverage of WES was 65× [range 40–94×] was significantly different among the EGFR TKI treatment group which might impact the detection number of low allelic fraction mutation. However, this sequencing coverage depth is still enough to define a significant pathway and genomic alteration that correlates with de novo resistance. Fourth, the copy number threshold to define amplification was adjusted (>0.3), more precisely than the pipeline recommendation. The gene-level segment integer copy number was parallel conducted using PureCN. The algorithm was previously shown good concordance with absolute copy number by targeted NGS- Foundation Medicine platform<sup>54</sup>. However, currently hybrid capture-based NGS has diverse thresholds for amplification. The threshold used in FoundationOne® Heme for identifying a copy number amplification is 5 for *ERBB2* and 6 for all other genes while the threshold used in Origimed is 6 for all. Lastly, we didn't assess prognostic significance for non-significant genes such as *AURKB* (1.5%) and *RBM10* (6%), even reported the prognostic significance associated with EGFR TKI. Four of 65 discrepancies of EGFR mutation results between WES and Cobas® mutation testing were found. Integrative Genomics Viewer analysis (IGV) on the bam file was performed; all specimens had EGFR exon 19 deletion reads less than 15 which were removed by our algorithm.

## Materials and methods

**Study population.** All methods were carried out in accordance with the declarations of Helsinki. The Institutional Review Board (IRB), Faculty of Medicine, Chulalongkorn University approved the study protocol (IRB 298/60). Written Informed consent was waived from individual study participants according to the ethics committee/IRB, Faculty of Medicine, Chulalongkorn University policy for retrospective study. The permission to conduct the study was provided by the director of the hospital. We retrospectively analyzed patients with pathologically confirmed recurrence or metastatic NSCLC diagnosed between 2011 and 2018 at King Chulalongkorn Memorial Hospital (KCMH). EGFR mutation testing was determined by cobas® EGFR Mutation Test v2 kit. Patients with NSCLC harboring activating EGFR mutations who received 1st or 2nd generation EGFR TKIs were included in our study, excluding osimertinib according to limited participants (n = 9) (Fig. 1). All patients were assessed for tumor responses and followed up every two to three months as the standard protocol of our institution. Objective response rate (ORR) was determined according to the Response Evaluation Criteria in Solid Tumors version 1.1 (RECIST v1.1) and classified as a progressive disease (PD), complete response (CR), partial response (PR), or stable disease (SD). Patients were categorized into three groups based on responsiveness to EGFR TKI treatment: (1) those with de novo EGFR TKI resistance who were defined as the best response were PD or SD less than 3 months while receiving EGFR TKI<sup>16</sup> (de novo resistance); (2) those who developed acquired resistance to EGFR TKIs according to the proposed criteria by Jackman<sup>17</sup> (Intermediate responder) and (3) those treated with EGFR TKIs for at least 2 years<sup>18</sup> (Long-term responder). Independent radiologist blind to molecular characteristic had reviewed imaging responses of 65 patients who obtained available tissue for WES.

**Exome sequencing analysis.** Genomic DNA was extracted from paraffin-embedded tissue, using Qiagen FFPE DNA extraction kits following manufacturer protocol. We used leftover extracted genomic DNA after cobas® EGFR Mutation Test as part of standard testing in advanced stage disease. After performing quality control (QC), qualified samples were proceeded to library construction. The genomic library was constructed with SureSelectXT V6 + UTR library prep kit (Illumina, San Diego, CA, USA) and was sequenced using NovoSeq to generate 150 bp paired-end reads at Macrogen Inc. (Seoul, Korea). The analytical pipeline of “cohort-normal” which showed high concordance rate ( $R^2$  0.99) to “match-normal” workflow in 10 significant pathways (206 genes) was explored in our study. The analysis of “cohort-normal” workflow was compared with our “match-normal” workflow, using 65 pair tumor-normal fresh tissue WES from resectable lung cancer patients who had received surgical procedures at The King Chulalongkorn Memorial Hospital. Written Informed consent was obtained in all resectable lung cancer participants. We selected those 65 specimens as retrospective manner based on EGFR TKI response with an adequate amount of specimen for WES, enriched in de novo EGFR TKI resistance. Ninety-eight percent of the second cohort WES had sensitizing mutation; composed of 55% EGFR exon 19 deletion, 43% exon 19 L858R, and one patient with exon 21 L861Q mutation. Sixty-five WES advanced stage NSCLC were categorized to 11 de novo resistance, 43 intermediate responders and 11 long-term responders. Pooled normal in the “cohort-normal” pipeline was obtained from either 65 normal lung tissue or leucocytes from the first exploratory cohort. Preprocessing steps and filtering for variant and copy number alteration (CNAs) discovery are described in the supplementary information.

**Statistical analysis.** The correlation of all categorical variables was analyzed using Kruskal–Wallis test. Significant correlation of two categorical variables was analyzed using two-sided Fisher's exact test or Chi-squared test for *p*-value calculations, while correlation of two continuous variables was conducted using Wilcoxon rank-sum test. FDR *p*-values were calculated by Benjamini–Hochberg Method from all correlation *p*-values in this cohort. Progression-free survival (PFS) was calculated from the first day of treatment with EGFR TKI to disease

progression or death. Overall survival (OS) was calculated from the date of diagnosis of recurrent or metastatic disease until the date of death. Patients were censored on June 30, 2020, if alive. Survival analysis was performed with a Kaplan–Meier analysis and log-rank test. Multivariate analysis was performed by binary logistic regression or Cox’s proportional hazards regression model when appropriate. The level of statistical significance was determined as a  $p$ -value less than 0.05. All statistical analyses were conducted using SPSS 23.0 (SPSS Inc, Chicago, Illinois, USA) and R package version 3.6.3.

Received: 4 September 2021; Accepted: 17 January 2022

Published online: 09 February 2022

## References

- Shi, Y. *et al.* A prospective, molecular epidemiology study of EGFR mutations in Asian patients with advanced non-small-cell lung cancer of adenocarcinoma histology (PIONEER). *J. Thorac. Oncol.* **9**, 154–162. <https://doi.org/10.1097/JTO.000000000000033> (2014).
- Zhang, Y. L. *et al.* The prevalence of EGFR mutation in patients with non-small cell lung cancer: A systematic review and meta-analysis. *Oncotarget* **7**, 78985–78993. <https://doi.org/10.18632/oncotarget.12587> (2016).
- Sriuranpong, V. *et al.* High frequency of mutation of epidermal growth factor receptor in lung adenocarcinoma in Thailand. *Cancer Lett.* **239**, 292–297. <https://doi.org/10.1016/j.canlet.2005.08.029> (2006).
- Mok, T. S. *et al.* Gefitinib or carboplatin–paclitaxel in pulmonary adenocarcinoma. *N. Engl. J. Med.* **361**, 947–957. <https://doi.org/10.1056/NEJMoa0810699> (2009).
- Mitsudomi, T. *et al.* Gefitinib versus cisplatin plus docetaxel in patients with non-small-cell lung cancer harbouring mutations of the epidermal growth factor receptor (WJTOG3405): An open label, randomised phase 3 trial. *Lancet Oncol.* **11**, 121–128. [https://doi.org/10.1016/S1470-2045\(09\)70364-X](https://doi.org/10.1016/S1470-2045(09)70364-X) (2010).
- Maemondo, M. *et al.* Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. *N. Engl. J. Med.* **362**, 2380–2388. <https://doi.org/10.1056/NEJMoa0909530> (2010).
- Zhou, C. *et al.* Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): A multicentre, open-label, randomised, phase 3 study. *Lancet Oncol.* **12**, 735–742. [https://doi.org/10.1016/S1470-2045\(11\)70184-X](https://doi.org/10.1016/S1470-2045(11)70184-X) (2011).
- Rosell, R. *et al.* Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): A multicentre, open-label, randomised phase 3 trial. *Lancet Oncol.* **13**, 239–246. [https://doi.org/10.1016/S1470-2045\(11\)70393-X](https://doi.org/10.1016/S1470-2045(11)70393-X) (2012).
- Wu, Y. L. *et al.* First-line erlotinib versus gemcitabine/cisplatin in patients with advanced EGFR mutation-positive non-small-cell lung cancer: Analyses from the phase III, randomized, open-label ENSURE study. *Ann. Oncol.* **26**, 1883–1889. <https://doi.org/10.1093/annonc/mdv270> (2015).
- Sequist, L. V. *et al.* Phase III study of afatinib or cisplatin plus pemetrexed in patients with metastatic lung adenocarcinoma with EGFR mutations. *J. Clin. Oncol.* **31**, 3327–3334. <https://doi.org/10.1200/JCO.2012.44.2806> (2013).
- Wu, Y. L. *et al.* Afatinib versus cisplatin plus gemcitabine for first-line treatment of Asian patients with advanced non-small-cell lung cancer harbouring EGFR mutations (LUX-Lung 6): An open-label, randomised phase 3 trial. *Lancet Oncol.* **15**, 213–222. [https://doi.org/10.1016/S1470-2045\(13\)70604-1](https://doi.org/10.1016/S1470-2045(13)70604-1) (2014).
- Riely, G. J. *et al.* Clinical course of patients with non-small cell lung cancer and epidermal growth factor receptor exon 19 and exon 21 mutations treated with gefitinib or erlotinib. *Clin. Cancer Res.* **12**, 839–844. <https://doi.org/10.1158/1078-0432.CCR-05-1846> (2006).
- Jackman, D. M. *et al.* Exon 19 deletion mutations of epidermal growth factor receptor are associated with prolonged survival in non-small cell lung cancer patients treated with gefitinib or erlotinib. *Clin. Cancer Res.* **12**, 3908–3914. <https://doi.org/10.1158/1078-0432.CCR-06-0462> (2006).
- Wang, J., Wang, B., Chu, H. & Yao, Y. Intrinsic resistance to EGFR tyrosine kinase inhibitors in advanced non-small-cell lung cancer with activating EGFR mutations. *Onco Targets Ther.* **9**, 3711–3726. <https://doi.org/10.2147/OTT.S106399> (2016).
- Stewart, E. L., Tan, S. Z., Liu, G. & Tsao, M. S. Known and putative mechanisms of resistance to EGFR targeted therapies in NSCLC patients with EGFR mutations—a review. *Transl. Lung Cancer Res.* **4**, 67–81. <https://doi.org/10.3978/j.issn.2218-6751.2014.11.06> (2015).
- Lee, J. K. *et al.* Primary resistance to epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (TKIs) in patients with non-small-cell lung cancer harboring TKI-sensitive EGFR mutations: an exploratory study. *Ann. Oncol.* **24**, 2080–2087. <https://doi.org/10.1093/annonc/mdt127> (2013).
- Jackman, D. *et al.* Clinical definition of acquired resistance to epidermal growth factor receptor tyrosine kinase inhibitors in non-small-cell lung cancer. *J. Clin. Oncol.* **28**, 357–360. <https://doi.org/10.1200/JCO.2009.24.7049> (2010).
- Gautschi, O. *et al.* Mutation profiling of lung cancers with long-term response to gefitinib therapy. *Oncol. Res. Treat.* **38**, 560–569. <https://doi.org/10.1159/000441367> (2015).
- Chen, J. *et al.* Genomic landscape of lung adenocarcinoma in East Asians. *Nat. Genet.* **52**, 177–186. <https://doi.org/10.1038/s41588-019-0569-6> (2020).
- Oh, S. *et al.* Reliable analysis of clinical tumor-only whole-exome sequencing data. *JCO Clin. Cancer Inform.* **4**, 321–335. <https://doi.org/10.1200/CCL19.00130> (2020).
- Sanchez-Vega, F. *et al.* Oncogenic signaling pathways in the cancer genome atlas. *Cell* **173**, 321–337. <https://doi.org/10.1016/j.cell.2018.03.035> (2018).
- Gini, B., Thomas, N. & Blakely, C. M. Impact of concurrent genomic alterations in epidermal growth factor receptor (EGFR)-mutated lung cancer. *J. Thorac. Dis.* **12**, 2883–2895. <https://doi.org/10.21037/jtd.2020.03.78> (2020).
- Cao, J. *et al.* An accurate and comprehensive clinical sequencing assay for cancer targeted and immunotherapies. *Oncologist* **24**, e1294–e1302. <https://doi.org/10.1634/theoncologist.2019-0236> (2019).
- Park, S. *et al.* Discordance of molecular biomarkers associated with epidermal growth factor receptor pathway between primary tumors and lymph node metastasis in non-small cell lung cancer. *J. Thorac. Oncol.* **4**, 809–815. <https://doi.org/10.1097/JTO.0b013e3181a94af4> (2009).
- Monaco, S. E. *et al.* A comparison of EGFR and KRAS status in primary lung carcinoma and matched metastases. *Hum. Pathol.* **41**, 94–102. <https://doi.org/10.1016/j.humpath.2009.06.019> (2010).
- Hasegawa, Y. *et al.* The role of smoking status on the progression-free survival of non-small cell lung cancer patients harboring activating epidermal growth factor receptor (EGFR) mutations receiving first-line EGFR tyrosine kinase inhibitor versus platinum doublet chemotherapy: A meta-analysis of prospective randomized trials. *Oncologist* **20**, 307–315. <https://doi.org/10.1634/theoncologist.2014-0285> (2015).

27. Han, J. Y. *et al.* Comparison of targeted next-generation sequencing with conventional sequencing for predicting the responsiveness to epidermal growth factor receptor-tyrosine kinase inhibitor (EGFR-TKI) therapy in never-smokers with lung adenocarcinoma. *Lung Cancer* **85**, 161–167. <https://doi.org/10.1016/j.lungcan.2014.04.009> (2014).
28. Lim, S. M. *et al.* Targeted sequencing identifies genetic alterations that confer primary resistance to EGFR tyrosine kinase inhibitor (Korean Lung Cancer Consortium). *Oncotarget* **7**, 36311–36320. <https://doi.org/10.18632/oncotarget.8904> (2016).
29. Sos, M. L. *et al.* PTEN loss contributes to erlotinib resistance in EGFR-mutant lung cancer by activation of Akt and EGFR. *Cancer Res.* **69**, 3256–3261. <https://doi.org/10.1158/0008-5472.CAN-08-4055> (2009).
30. Engelman, J. A. *et al.* MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* **316**, 1039–1043. <https://doi.org/10.1126/science.1141478> (2007).
31. Cappuzzo, F. *et al.* MET increased gene copy number and primary resistance to gefitinib therapy in non-small-cell lung cancer patients. *Ann. Oncol.* **20**, 298–304. <https://doi.org/10.1093/annonc/mdn635> (2009).
32. Benedettini, E. *et al.* Met activation in non-small cell lung cancer is associated with de novo resistance to EGFR inhibitors and the development of brain metastasis. *Am. J. Pathol.* **177**, 415–423. <https://doi.org/10.2353/ajpath.2010.090863> (2010).
33. Qin, K., Hou, H., Liang, Y. & Zhang, X. Prognostic value of TP53 concurrent mutations for EGFR-TKIs and ALK-TKIs based targeted therapy in advanced non-small cell lung cancer: a meta-analysis. *BMC Cancer* **20**, 328. <https://doi.org/10.1186/s12885-020-06805-5> (2020).
34. Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339**, 1546–1558. <https://doi.org/10.1126/science.1235122> (2013).
35. Vogelstein, B. & Kinzler, K. W. Cancer genes and the pathways they control. *Nat. Med.* **10**, 789–799. <https://doi.org/10.1038/nm1087> (2004).
36. Lee, T. G. *et al.* Altered expression of cellular proliferation, apoptosis and the cell cycle-related genes in lung cancer cells with acquired resistance to EGFR tyrosine kinase inhibitors. *Oncol. Lett.* **14**, 2191–2197. <https://doi.org/10.3892/ol.2017.6428> (2017).
37. Blakely, C. M. *et al.* Evolution and clinical impact of co-occurring genetic alterations in advanced-stage EGFR-mutant lung cancers. *Nat. Genet.* **49**, 1693–1704. <https://doi.org/10.1038/ng.3990> (2017).
38. Wang, Q., Yang, S., Wang, K. & Sun, S. Y. MET inhibitors for targeted therapy of EGFR TKI-resistant lung cancer. *J. Hematol. Oncol.* **12**, 63. <https://doi.org/10.1186/s13045-019-0759-9> (2019).
39. Ramalingam, S. S. *et al.* LBA50—Mechanisms of acquired resistance to first-line osimertinib: Preliminary data from the phase III FLAURA study. *Ann. Oncol.* **29**, 740. <https://doi.org/10.1093/annonc/mdy424.063> (2018).
40. Shi, P. *et al.* Met gene amplification and protein hyperactivation is a mechanism of resistance to both first and third generation EGFR inhibitors in lung cancer treatment. *Cancer Lett.* **380**, 494–504. <https://doi.org/10.1016/j.canlet.2016.07.021> (2016).
41. Ortiz-Cuaran, S. *et al.* Heterogeneous mechanisms of primary and acquired resistance to third-generation EGFR inhibitors. *Clin. Cancer Res.* **22**, 4837–4847. <https://doi.org/10.1158/1078-0432.CCR-15-1915> (2016).
42. Leonetti, A. *et al.* Resistance mechanisms to osimertinib in EGFR-mutated non-small cell lung cancer. *Br. J. Cancer* **121**, 725–737. <https://doi.org/10.1038/s41416-019-0573-8> (2019).
43. Kim, Y. *et al.* Concurrent genetic alterations predict the progression to target therapy in EGFR-mutated advanced NSCLC. *J. Thorac. Oncol.* **14**, 193–202. <https://doi.org/10.1016/j.jtho.2018.10.150> (2019).
44. Bhateja, P. *et al.* Retinoblastoma mutation predicts poor outcomes in advanced non small cell lung cancer. *Cancer Med.* **8**, 1459–1466. <https://doi.org/10.1002/cam4.2023> (2019).
45. Anders, L. *et al.* A systematic screen for CDK4/6 substrates links FOXM1 phosphorylation to senescence suppression in cancer cells. *Cancer Cell* **20**, 620–634. <https://doi.org/10.1016/j.ccr.2011.10.001> (2011).
46. Wang, H. *et al.* The metabolic function of cyclin D3-CDK6 kinase in cancer cell survival. *Nature* **546**, 426–430. <https://doi.org/10.1038/nature22797> (2017).
47. Lagarrigue, S. *et al.* CDK4 is an essential insulin effector in adipocytes. *J. Clin. Invest.* **126**, 335–348. <https://doi.org/10.1172/JCI81480> (2016).
48. Deng, J. *et al.* CDK4/6 inhibition augments antitumor immunity by enhancing T-cell activation. *Cancer Discov.* **8**, 216–233. <https://doi.org/10.1158/2159-8290.CD-17-0915> (2018).
49. La Monica, S. *et al.* Efficacy of the CDK4/6 dual inhibitor abemaciclib in EGFR-mutated NSCLC cell lines with different resistance mechanisms to osimertinib. *Cancers* <https://doi.org/10.3390/cancers13010006> (2020).
50. Piotrowska, Z. *et al.* Landscape of acquired resistance to osimertinib in EGFR-mutant NSCLC and clinical validation of combined EGFR and RET inhibition with osimertinib and BLU-667 for acquired RET fusion. *Cancer Discov.* **8**, 1529–1539. <https://doi.org/10.1158/2159-8290.CD-18-1022> (2018).
51. Zeng, L., Yang, N. & Zhang, Y. GPCR-ROS1 rearrangement as an acquired resistance mechanism to osimertinib and responding to crizotinib combined treatments in lung adenocarcinoma. *J. Thorac. Oncol.* **13**, e114–e116. <https://doi.org/10.1016/j.jtho.2018.02.005> (2018).
52. Schrock, A. B. *et al.* Receptor tyrosine kinase fusions and BRAF kinase fusions are rare but actionable resistance mechanisms to EGFR tyrosine kinase inhibitors. *J. Thorac. Oncol.* **13**, 1312–1323. <https://doi.org/10.1016/j.jtho.2018.05.027> (2018).
53. Offin, M. *et al.* Acquired ALK and RET gene fusions as mechanisms of resistance to osimertinib in EGFR-mutant lung cancers. *JCO Precis. Oncol.* <https://doi.org/10.1200/PO.18.00126> (2018).
54. Riester, M. *et al.* PureCN: Copy number calling and SNV classification using targeted short read sequencing. *Source Code Biol. Med.* **11**, 13. <https://doi.org/10.1186/s13029-016-0060-z> (2016).

## Acknowledgements

This research was supported by the Ratchadapiseksompotch Fund, Faculty of Medicine, Chulalongkorn University (Grant number RA60/121) to CV, Chulalongkorn Academic Advancement into Its 2nd Century (CUAAC) Project to VS and Health Systems Research Institute (Thailand) (Grant number 64-106) to CV. Biospecimen collection was supported by Biobank, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand. The authors would like to thank Dr. Harit Suwanrusme for critical review of the manuscript.

## Author contributions

P.S. contributed to the conception, investigation, analysis and drafted the manuscript. C.T. contributed to data acquisition. K.K., N.L. and N.P. contributed to provide resources. N.P. contributed to conception and data acquisition. P.C. was contributed to data acquisition. S.S. and V.S. contributed to provide consultation. C.A. contributed to data acquisition and analysis of the data. W.C. contributed to imaging review. C.V. contributed to the design of the research, conception, analysis of the data, funding acquisition and drafted the manuscript. All authors critically revised the manuscript, agree to be fully accountable for ensuring the integrity and accuracy of the work, and approved the final manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-06239-y>.

**Correspondence** and requests for materials should be addressed to C.V.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022