EGFR DNA Methylation Correlates With EGFR Expression, Immune Cell Infiltration, and Overall Survival in Lung Adenocarcinoma

Zhanyu Xu¹, Fanglu Qin², Liqiang Yuan¹, Jiangbo Wei¹, Yu Sun¹, Junqi Qin¹, Kun Deng¹, Tiaozhan Zheng¹ and Shikang Li¹*

¹ Department of Thoracic and Cardiovascular Surgery, The First Affiliated Hospital of Guangxi Medical University, Nanning, China, ² School of Information and Management, Guangxi Medical University, Nanning, China

Background: The epidermal growth factor receptor (EGFR) is a primary target of molecular targeted therapy for lung adenocarcinoma (LUAD). The mechanisms that lead to epigenetic abnormalities of EGFR in LUAD are still unclear. The purpose of our study was to evaluate the abnormal methylation of EGFR CpG sites as potential biomarkers for LUAD.

Methods: To assess the differentially methylation CpG sites of EGFR in LUAD, we used an integrative study of Illumina HumanMethylation450K and RNA-seq data from The Cancer Genome Atlas (TCGA). We evaluated and compared EGFR multiple-omics data to explore the role of CpG sites located in EGFR promoter regions and gene body regions and the association with transcripts, protein expression levels, mutations, and somatic copy number variation. We calculated the correlation coefficients between CpG sites of EGFR and immune infiltration fraction (by MCPcounter and ESTIMATE) and immune-related pathways in LUAD. Finally, we validated the differential methylation of clinically and prognostically relevant CpG sites using quantitative methylation-specific PCR (qMSP).

Results: We found that the methylation level of many EGFR CpGs in the promoter region was negatively correlated with the transcription level, protein expression, and SCNV, while the methylation at the gene body region was positively correlated with these features. The methylation level of EGFR CpGs in the promoter region was positively correlated with the level of immune infiltration and IFN-γ signature, while the opposite was found for methylation of the gene body region. The qMSP results showed that cg02316066 had a high methylation level, while cg02166842 had a low methylation level in LUAD. There was a high degree of co-methylation between cg02316066 and cg03046247.

Conclusion: Our data indicate that EGFR is an epigenetic regulator in LUAD acting through DNA methylation. Our research provides a theoretical basis for the further detection of EGFR DNA methylation as a predictive biomarker for LUAD survival and immunotherapy.

Keywords: lung adenocarcinoma, EGFR, DNA methylation, tumor biomarkers, tumor-infiltrating
INTRODUCTION

Lung cancer is the primary cause of cancer-related death worldwide (1). The most prominent pathological subtype of lung cancer is lung adenocarcinoma (LUAD), which accounts for about 45 percent of lung cancer cases (2). The five-years overall survival rate for patients with advanced lung cancer is less than 20% (3). Genetic analyses have revealed driver genes in LUAD and have changed the treatment paradigm (4). In Asia, epidermal growth factor receptor (EGFR) mutations account for 51.4% of advanced LUAD driver mutations, while it accounts for 15 to 22% of advanced LUAD driver mutations in non-Asian areas (5, 6). EGFR tyrosine kinase inhibitors (TKIs) are typically used to treat patients with EGFR-mutant LUAD (7, 8). The high heterogeneity of this type of cancer, on the other hand, restricts the survival advantage of patients undergoing EGFR-targeted treatment, indicating the need for more research into new prognosis-related molecular mechanisms. The molecular mechanism by which EGFR regulates LUAD through DNA methylation has yet to be completely elucidated.

EGFR is a receptor tyrosine kinase (TK) that dimerizes in response to ligand stimulation, resulting in the activation of intracellular TKs and autophosphorylation of multiple tyrosine residues, which triggers a sequence of downstream signaling cascades (9, 10). The Ras/MAPK and PI3K/PKB signaling pathways are two of the most studied EGFR pathways, both of which have a well-established role in tumor development, survival, and progression (11). The most studied epigenetic mechanism is DNA methylation, which is linked to cell division, immune regulation, and X chromosome inactivation (12). Methylation of gene promoter regions is often linked to transcriptional silencing, while methylation of gene bodies has the opposite effect. Epigenetic dysregulation has been linked to the early stages of oncogenic transformation in a variety of solid tumors, and it can be used as a biomarker for early detection, systemic sampling, and prognosis in a variety of human cancers (13). Indeed, Haijing Liu et al. (14) reported the potential link between EGFR alterations at the multi-omics levels and clinical prognosis by pan-cancer analysis, but the relationship between DNA methylation of EGFR in LUAD and immune infiltration has not been reported. For patients with LUAD who do not benefit from targeted therapy, the discovery of DNA methylation-related immune landscapes has important implications for the molecular mechanisms of immunotherapy.

Using the LUAD dataset from The Cancer Genome Atlas (TCGA), we performed a comprehensive multi-omics data assessment of EGFR-annotated CpGs. We investigated whether EGFR CpG methylation sites correlated with EGFR gene expression, protein levels and overall survival (OS) time. We further explored the relationship between somatic copy number variation (SCNV) and DNA methylation of EGFR, and the link between EGFR CpG methylation sites and LUAD immunological infiltration cells and immune-related pathways.

MATERIALS AND METHODS

EGFR CpGs and TCGA Data Analysis

The Illumina 450K TCGA dataset was used to extract methylation and expression data for all EGFR CpGs. We investigated 49 CpGs in the promoter and gene body regions of EGFR. Figure 1 depicts the EGFR genome arrangement (Figure 1A) and the relative locations of all EGFR CpGs genomes (Figure 1B). Clinical information is integrated with data from the GDC data portal (https://portal.gdc.cancer.gov) (15). These data included 535 LUAD patients, as well as information on EGFR status and somatic copy number variation (SCNV). Methylation data came in the form of beta-values, while expression data came in the form of TPM (Transcripts Per kilobase Million)-normalized read counts. cBioPortal (http://www.cbioportal.org/study?id=luad_tcgaw) was used to retrieve RPPA (Reverse phase protein array)-based protein expression data (16).

Survival Analysis

For all available LUAD samples, Kaplan-Meier survival analysis curves (17) for the 49 EGFR CpG were plotted, with a P-value of 0.05 used as a statistical threshold, according to the group with high or low methylation. We did the above survival analysis curves in the EGFR wild-type group, EGFR mutation group and EGFR wild-type and PD-L1 high expression group.

Assessment of Immune Cells Infiltration, ESTIMATE Scores

We employed the R package ESTIMATE (18) to investigate the immune invasion of LUAD samples. After that, to obtain a more detailed picture of immune cell-types and other stromal cells infiltration, R package MCPCounter was used (19). MCPCounter utilizes the scoring data for individual tumor specimens (20). We calculated the Pearson correlation between the β value of the CpG site on EGFR and the score of immune infiltration.

Sample Collection

20 paired LUAD and non-cancerous lung tissue samples were obtained from the First Affiliated Hospital of Guangxi Medical University from September 2019 to February 2020, and stored at -80°C. LUAD diagnosis was confirmed by two independent pathologists. This study was approved by the ethics committee of the First Affiliated Hospital of Guangxi Medical University.

DNA Extraction, DNA Sodium Bisulfite Conversion, and Quantitative Methylation Specific PCR

Genomic DNA was isolated from individual specimens using a CTAB DNA extraction (21). Nanodrop 2000 spectrophotometer (Thermofisher, USA) was used to detect the concentration and purity of the DNA extraction, and nucleic acid gel electrophoresis was used to detect DNA integrity (22). Samples were then bisulfite-converted using the Epitect Fast DNA Bisulfite Kit
(Qiagen; 59824), according to the manufacturer protocol. The purified products were quantitated using a Qubit ssDNA Assay kit (Thermo, Q10212). Primer Premier 6.0 software (Premier, Canada) was used to design the primer sequences to target CpG sites. Fully methylated genomic DNA after bisulfite treatment and normal genomic DNA (not transformed with bisulfite) were used as templates. Uncalibrated methylation levels, roughly equivalent to percent methylation, were calculated using cycle threshold (CT) values obtained from probes that specifically bind to methylated (CT_methylated) and unmethylated (CT_unmethylated) DNA, respectively (methylation [%] = 100%/(1 + 2^{CT_methylated/CT_unmethylated}). The primers were utilized as shown in Table 1. QMSP was performed using an Applied Biosystems 7900HT Fast Real-Time PCR system (Waltham, Massachusetts, USA) with the following temperature profile: 5 minutes at 95°C, followed by 40 cycles of 15 seconds at 95°C, 30 seconds at 60°C, and 60 seconds at 60°C.

**Statistical Analysis**

Pearson correlation coefficients were used to assess correlations between EGFR mRNA expression, protein expression, immune score, and all individual beta values of EGFR in the TCGA dataset. Wilcoxon rank sum test with continuity correction was used to assess differential methylation. Results were deemed significant if p<0.05.

**RESULTS**

**Differential Methylation Analysis of EGFR and Its CpG Sites**

We sought to explore whether variations in DNA methylation were linked to EGFR expression abnormalities. The chiproterminal official website analysis showed that EGFR hypermethylation is inversely correlated with mRNA ($r^2 = -0.38, P = 1.14e-16$) (Figure 1C) and protein ($r^2 = -0.39, P = 5.76e-13$) (Figure 1D) overexpression in LUAD of TCGA. We next analyzed the relationship between the methylation levels of the EGFR promoter and the clinicopathological parameters of LUAD patients by UALCAN. EGFR were significantly hypermethylated in LUAD tissues when compared with normal lung tissues ($P = 3.38e-3$) (Figure 1E), and the mRNA levels of EGFR in LUAD were remarkably lower than those in normal lung tissues ($P = 1.41e-4$) (Figure 1F).
In the same cohort of LUAD patients from the TCGA database we collected methylation data from the Infinium HumanMethylation450 BeadChip for 49 CpG sites of EGFR (Table 2 and Supplemental Table 1). Six CpG sites were located in promoter regions (cg16751451, cg07311521, cg03860890, cg22396409, cg10949960) and 43 were in the gene body or in the 3' UTR regions. There were 34 of 49 CpG sites that were differentially methylated between LUAD tissues and control groups (P < 0.05). Five CpG sites in promoter regions and 17 CpG sites in the gene body had a significantly higher percentage of methylation in LUAD when compared to normal lung tissues. Meanwhile, cg16751451 in promoter regions and 11 CpG sites in the gene body or 3' UTR regions had a significantly lower percentage of methylation in LUAD compared to normal lung tissues.

LUAD with EGFR mutations is a subtype of LUAD with a particular molecular mechanism and selective treatment (23). We analyzed EGFR methylation changes of LUAD in the EGFR mutations group and EGFR wild-type (non-mutated) group. We found that in the EGFR mutation group, EGFR showed a significantly higher percentage of methylation in LUAD compared to normal lung tissues and control groups (P < 0.05). Five CpG sites in promoter regions and 17 CpG sites in the gene body had a significantly higher percentage of methylation in LUAD when compared to normal lung tissues. Meanwhile, cg16751451 in promoter regions and 11 CpG sites in the gene body or 3' UTR regions had a significantly lower percentage of methylation in LUAD compared to normal lung tissues.

We next performed a survival analysis in the EGFR-mutant and EGFR-wild subsets separately. In the EGFR mutation group, the hypermethylation of cg01461514, cg26277197 and cg25311271 was associated with a good prognosis of LUAD (p<0.05), while the hypermethylation of cg21901928, cg22427313, cg02316066, cg26055062, cg10002850 and cg03046247 was associated with a poor prognosis (p<0.05) (Figure 2B). In patients with wild type EGFR, high levels of PDL1 expression affected the prognosis of immunotherapy (24).

Apart from EGFR mutations, KRAS mutations were the most common mutations in LUAD (25), and we discovered that hypermethylation of cg26055062 and cg04625338 predicted a good prognosis in LUAD patients with KRSA mutations (Figure 3A). Hypermethylation of cg18809076 and cg25311271 in the KRAS wild-type group predicted good prognosis, but ch.7:1264585R had the opposite effect (Figure 3B). Patients with LUAD might have a variety of molecular features depending on their smoking history (26). We explored the relationship between EGFR methylation and prognosis in patients with LUAD in the smoking and non-smoking groups. Hypermethylation of cg16751451 and cg27637738 suggested a poor prognosis for LUAD in the smoking group (Figure 3C). In the nonsmoking group, hypermethylation of cg04625338 and cg05064645 indicated a favorable outcome, but cg02316066, cg05898452, cg06052090, cg21808635, cg16751451, and cg16589260 had the opposite effect (Figure 3D).

### Table 1

| Name               | Sequence (5'-3')       | Types                      | Products (bp) |
|--------------------|------------------------|----------------------------|---------------|
| cg02316066-M       | TGTGGGGTTACGCGTTAAGTTTC| Forward Primer             | 170           |
| cg02316066-M       | TCTACCACTATAAATATCATACAT| Reverse Primer             | 170           |
| cg02316066-U       | TGTGGGGTTACGCGTTAAGTTT | Forward Primer             | 170           |
| cg02316066-U       | TCTACCACTATAAATATCATACAT| Reverse Primer             | 170           |
| cg03046247-M       | TGGAAATAGTAAATCGGAGTGGA | Forward Primer             | 228           |
| cg03046247-M       | AACTACGCTATTTTTAAAAACCACG| Reverse Primer            | 228           |
| cg03046247-U       | TGGAAATAGTAAATCGGAGTGGA| Forward Primer             | 228           |
| cg03046247-U       | AAAACTACACATTTTAAAAACCACA| Reverse Primer            | 228           |
| cg02166842-M       | GAGTGAGTGTTGTTAATGAGTGATG| Forward Primer            | 170           |
| cg02166842-M       | AOCCTCTCAAATATATATTTTACAG | Reverse Primer      | 170           |
| cg02166842-U       | GAGTGAGTGTTGTTAATGAGTGATG| Forward Primer            | 170           |
| cg02166842-U       | AOCCTCTCAAATATATATTTTACAG| Reverse Primer            | 170           |

M, methylated; U, unmethylated.

### OS-Related CpG Sites

To explore the prognostic value of 49 CpGs of EGFR in LUAD, we constructed survival curves to evaluate the association between CpGs and OS with the Kaplan-Meier method. A total of 10 CpG sites were significantly associated with the OS of LUAD patients (Figure 2A). Except for cg05064645, the hypermethylation of cg27637738, cg16751451, cg02316066, cg22396409, cg03046247, cg02166842, cg21901928, cg07311521 and cg06052090 CpG sites revealed poor prognosis of LUAD patients (p<0.05).

### EGFR Expression Is Correlated With DNA Methylation

Individual CpG methylation was studied in relation to EGFR mRNA and protein expression (Figure 4A). Of the 49 CpG sites examined in LUAD tissue, 44 had a strong association with EGFR mRNA expression. The methylation levels of all CpG sites...
in promoter regions were shown to be inversely correlated to EGFR mRNA levels. The methylation level of CpG sites in 27 of 43 CpGs within the gene body region was positively correlated with the mRNA level of EGFR, and 11/43 showed a significant negative correlation. Cg03046247, cg08428266 and cg20062492 hypermethylation was significantly related to the high expression of EGFR mRNA, and cg10002850 hypomethylation was significantly related to the high expression of EGFR mRNA.

Then we investigated the connection between CpG methylation and EGFR protein expression, and the findings matched the EGFR mRNA association described previously. Cg02316066 hypermethylation and cg01461514 hypomethylation were significantly related to the high expression of EGFR protein.

Next, we explored the correlation between the mRNA and protein expression levels of EGFR and EGFR CpG sites in EGFR and KRAS mutation status or not (Figures 4B, C).
FIGURE 2 | KM curves of EGFR CpG sites. (A) Kaplan-Meier analysis regarding overall survival in patients with LUAD stratified according to EGFR methylation CpG sites. (B, C) Kaplan-Meier analysis of the overall survival rate of LUAD patients based on the EGFR mutation group and the EGFR wild-type PDL1 high expression group.
FIGURE 3 | KM curves of EGFR CpG sites. Kaplan-Meier analysis of the overall survival rate of LUAD patients based on the KRAS mutation group (A), KRAS wild-type group (B), smoking group (C) and the non-smoking group (D).
Overall, similar to the above results, the β values of CpG sites located in the EGFR promoter region were negatively correlated with the mRNA and protein levels of EGFR in both the EGFR mutant and KRAS mutant groups, while the CpG sites of the gene body showed the opposite effect. Interestingly, we found a higher correlation, both positive and negative, in the EGFR mutant group compared to the EGFR wild-type group. In contrast, this correlation was lower in the KRAS mutation group than in the KRAS wild-type group.

**EGFR Methylation and Expression Are Associated With Immune Cells Infiltration**

Immune cells from a variety of species are known to infiltrate the tumor microenvironment (27). We explored the association between EGFR methylation levels and the infiltration levels of 8 immune cells and 2 stromal cells of MCPcounter. Based on the median EGFR integrated methylation level, we divided the LUAD samples into hypermethylated and hypomethylated EGFR groups. The results showed that the EGFR hypermethylation group was associated with increased infiltration of T cells, CD8 T cells, cytotoxic lymphocytes, B lineage, NK cells, monocytic lineage, and fibroblasts (Figure 5A). Moreover, the immune score, stromal score and estimate score of ESTIMATE were higher in the EGFR hypermethylated group than those of the hypomethylated group (Figure 5B). We found a positive association between EGFR promoter hypermethylation and the infiltration of T cells, CD8 T cells, cytotoxic lymphocytes, B lineage, NK cells, monocytic lineage, endothelial cells and fibroblasts, while EGFR body hypomethylation had lower infiltration of the above immune cells (Figure 5C). To further examine this scenario, we found that EGFR promoter hypermethylation had higher immune scores (p<0.001), a marker of total immune infiltration (Figure 5D). The association between EGFR methylation and the IFN-signature was also investigated. Increased mRNAs of the main IFN-signature genes (IFNG, STAT1, STAT2, JAK2) were found to be linked to extensive promoter hypermethylation and gene body hypomethylation (Figure 5E).

**Validation of CpG Site Methylation by qMSP in External Cohorts**

We evaluated the relationship between CpG sites and clinical information, such as T staging, N staging and cancer states. As shown in Figure 6, the increase in the methylation β value of cg02166842 was significantly correlated with the T stage (T1 vs T2, T1 vs T3) (p=0.027, p=0.025 respectively) (Figure 6A), N stage (N0 vs N2) (p=0.035) and with tumor status (p=0.032) (Figure 6C). The methylation β value of cg02316066 was positively correlated with the T stage (T1 vs T3, T1 vs T4, T2 vs T3) (p=0.0049, p=0.019, p=0.032 respectively) (Figure 6D), while the methylation β value of cg03046247 was positively correlated with the T stage (T1 vs T2, T1 vs T3, T1 vs T4, T2 vs T3, T2 vs T4) (p=0.026, p=0.0011, p=0.0038, p=0.049, p=0.028 respectively) (Figure 6E) and
To determine whether there was a difference in the methylation levels of cg02316066, cg03046247 and cg02166842 between LUAD tissue and adjacent tissues, we collected 20 pairs of tissue specimens. Consistent with the TCGA database results, cg02316066 showed hypermethylation levels and cg02166842 showed hypomethylation in LUAD (Figures 6G, H). Ten OS-related CpG sites showed high correlation coefficients towards each other indicating a high degree of co-methylation (Figure 6I). We noted that cg02316066 and cg03046247 were strongly associated with multiple types of clinical profiles and LUAD prognosis, and there was also a high degree of co-methylation between cg02316066 and cg03046247. In the TCGA cohort, the cg02316066 and cg03046247 Pearson correlation coefficient was 0.76 (p<0.001) (Figure 6J) and, consistent with this, in our validation cohort the correlation coefficient was 0.56 (p<0.001) (Figure 6K).

**DISCUSSION**

So far, EGFR is a consensual factor that promotes cancer progression and the development of EGFR-TKIs has dramatically changed the therapeutic landscape for patients with non-small cell lung cancer (28). However, the rapid occurrence of clinical drug resistance hinders patient survival (8). TKIs or monoclonal antibodies targeting EGFR can block the
infiltration of immunosuppressive cells and improve the antitumor response in NSCLCs, indicating that combining EGFR-targeted therapy with immune checkpoint inhibitors is a viable alternative for combination immunotherapy (29).

Epigenetic variations are being gradually investigated, and they are now changing the idea that malignant lesions depend entirely on genetic expressions to develop (30). DNA methylation, which is largely responsible for gene silencing and chromatin formation, is by far the most studied epigenetic regulatory mechanism (31). Methyl groups are covalently bound to cytosine during DNA methylation to generate 5-methylcytosine (5mC) (32). Furthermore, DNA methylation is chemically stable, can be tested separately, and has strong biomarker potential (33). Methylation quantitative measurement of small samples (microanatomical cells, biopsies) is often performed in the clinic. These are the advantages of using methylation as a biomarker (34, 35).

To establish the various levels of methylation sites of EGFR in LUAD, we collected RNA-seq results and Illumina HumanMethylation450K from TCGA. We meticulously investigated DNA methylation at the EGFR 49 CpG sites. We correlated EGFR methylation with transcription and protein expression in LUAD tissues. Our results indicate that DNA methylation in the promoter and gene body regions resulted in strong epigenetic regulation of EGFR. In the LUAD TCGA cohort, hypomethylation of the promoter region was negatively associated with increased mRNA and protein expression, while hypomethylation in the gene body was nearly always positively correlated with EGFR expression. This strong epigenetic regulation of EGFR is present not only in the different mutational states of EGFR but also in the KRAS mutant and wild-type groups, and interestingly, mutations in EGFR enhance this epigenetic regulation of EGFR, while mutations in KRAS attenuate this property. Some studies have revealed that EGFR and KRAS mutations are mutually exclusive in lung adenocarcinoma (36), the mechanisms of which need to be investigated in more depth. These methylation defects can eventually affect the clinical characteristics and prognosis of LUAD patients.

Mutations in the Furin-like and Pkinase-Tyr domains were shown to be predictive indicators of effective TKI therapy for
NSCLC (37–39), with slightly longer survival as compared to standard combination chemotherapy (40, 41). Different EGFR mutations have various benefits, and those that inhibit EGFR kinase activity can benefit from EGFR-targeted therapy (42, 43). In LUAD patients with an EGFR mutant phenotype, we found that most (19/26) CpG sites were hypomethylated and six of these were predictors for a good prognosis. There was a negative correlation between SCNV and DNA methylation at sites in CpG islands, and conversely, a positive correlation between sites in the CpG ocean of EGFR (26/31).

In LUAD patients with an EGFR mutant phenotype, we found that most (19/26) CpG sites were hypomethylated and six of these were predictors for a good prognosis. There was a negative correlation between SCNV and DNA methylation at sites in CpG islands, and conversely, a positive correlation between sites in the CpG ocean of EGFR (26/31).

T cells, CD8+ T cells, cytotoxic lymphocytes, B lineage, NK cells, monocyte lineage, and fibroblasts were found to be infiltrated more frequently in tissues that presented EGFR hypermethylation. Also, the immune score, stromal score and estimate score were higher in the EGFR hypermethylated group than those in the hypermethylated group. The non-inflammatory tumor microenvironment (TME) in EGFR-mutated NSCLCs is abundant in Treg cells and macrophages, with the latter releasing chemokines that attract more Treg cells in the inflammatory TME (44). EGFR-TKI therapy facilitates CD8+ T cell recruitment and prevents Treg cell infiltration in the TME in EGFR-mutated tumors in vivo (45). Therefore, we evaluated the relationship between EGFR methylation and an IFN-γ signature. Widespread promoter hypermethylation and body hypomethylation were strongly associated with increased IFN-γ signature. In vitro studies have shown that blocking EGFR with antibodies or kinase inhibitors facilitate the secretion of chemokines (CCL2, CCL5, and CXCL10) in HNSCC cells and keratinocytes when IFN and tumor necrosis factor (TNF) are stimulated (46).

In summary, our research shows that EGFR participates in the epigenetic regulation of LUAD through DNA methylation. DNA methylation of EGFR shows unique clinical characteristics and immunogenicity. Our research provides a theoretical basis for further assessment of EGFR DNA methylation, which can be used as a biomarker to predict the prognosis and immune mechanisms of LUAD.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

REFERENCES

1. Herbst RS, Morgensztern D, Boshoff C. The Biology and Management of Non-Small Cell Lung Cancer. Nature (2018) 553(7689):446–54. doi: 10.1038/nature25183
2. Zappa C, Moussa SA. Non-Small Cell Lung Cancer: Current Treatment and Future Advances. Transl Lung Cancer Res (2016) 5(3):288–300. doi: 10.21037/tlcr.2016.06.07
3. Kanodra NM, Silvestri GA, Tanner NT. Screening and Early Detection Efforts in Lung Cancer. Cancer (2015) 121(9):1347–56. doi: 10.1002/cncr.29222
4. Li T, Kung HJ, Mack PC, Gandara DR. Genotyping and Genomic Profiling of Non-Small-Cell Lung Cancer: Implications for Current and Future Therapies. J Clin Oncol (2013) 31(8):1039–49. doi: 10.1200/JCO.2012.45.3753
5. Cancer Genome Atlas Research Network. Comprehensive Molecular Profiling of Lung Adenocarcinoma. Nature (2014) 511(7511):543–50. doi: 10.1038/nature13385
6. Shi Y, Au JS, Thongprasert S, Srinivasan S, Tsai CM, Khoa MT, 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 (2014) 9(2):154–62. doi: 10.1097/JTO.0000000000000633
7. Remon J, Steuer CE, Ramalingam SS, Felip E, Osimertinib and Other Third-Generation EGFR TKI in EGFR-Mutant NSCLC Patients. Ann Oncol (2018) 29(suppl_1):i20–7. doi: 10.1093/annonc/mdy704
8. Wu SG, Shih JY. Management of Acquired Resistance to EGFR TKI-Targeted Therapy in Advanced Non-Small Cell Lung Cancer. Mol Cancer (2018) 17(1):38. doi: 10.1186/s12943-018-0777-1

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Ethics Committee of the First Affiliated Hospital of Guangxi Medical University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

ZX and FQ conceived and designed the experiments, performed the qMSP experiments, prepared figures and/or tables and approved the final draft. LY, KD, and TZ performed bioinformatics analysis, analyzed the data, authored or reviewed drafts of the paper and approved the final draft. YS and JQ prepared figures and/or tables and approved the final draft. SL conceived and designed the experiments, approved the final draft. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the National Natural Science Foundation of China (no. NSFC81660488) and the Guangxi Natural Science Foundation under grant no. 2017GXNSFAA198123.

ACKNOWLEDGMENTS

The authors thank the staff of the Department of Thoracic Surgery, The First Affiliated Hospital of Guangxi Medical University, and all participants and contributors to this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.691915/full#supplementary-material

Supplementary Table 1 | Annotation information of 49 CpG sites of EGFR.
80. Roskoski JR Jr. Small Molecule Inhibitors Targeting the EGFR/ErbB Family of Protein-Tyrosine Kinases in Human Cancers. Pharmacol Res (2019) 139:395–411. doi: 10.1016/j.phrs.2018.11.014
81. Cho I, Kim S, Du J, Meyerson M. Autophosphorylation of the Carboxyl-Terminal Domain Is Not Required for Oncogenic Transformation by Lung-Cancer Derived EGFR Mutants. Int J Cancer. (2018) 143(3):679–85. doi: 10.1002/ijc.31332
82. Siddiqui N, Sonenberg N. Signalling to eIF4E in Cancer. Biochem Soc Trans (2015) 43(5):763–72. doi: 10.1042/BST20150126
83. Meng H, Cao Y, Qin J, Song X, Zhang Q, Shi Y, et al. DNA Methylation, its Mediators and Genome Integrity. Int J Biol Sci (2015) 11(5):604–17. doi: 10.7150/ijbss.11218
84. Liu Y, Feng Y, Hou T, Lizaso A, Xu F, Xing P, et al. Investigation on the Potential of Circulating Tumor DNA Methylation Patterns as Prognostic Biomarkers for Lung Squamous Cell Carcinoma. Transl Lung Cancer Res (2020) 9(6):2356–66. doi: 10.21037/tlcr-20-1070
85. Liu H, Zhang B, Sun Z. Spectrum of EGFR Aberrations and Potential Clinical Implications: Insights From Integrative Pan-Cancer Analysis. Cancer Commun (Lond) (2020) 40(1):43–59. doi: 10.1007/s12943-018-0789-x
86. Zhang Z, Hernandez K, Savage J, Li S, Miller D, Agrawal S, et al. Uniform Genomic Data Analysis in the NCI Genomic Data Commons. Nat Commun (2021) 12(1):1226. doi: 10.1038/s41467-021-21254-9
87. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the Chipportal. Sci Signal (2013) 6(269):pl1. doi: 10.1126/sci signal.2004088
88. Lira RPC, Antunes-Foschi R, Rocha EM. Survival Analysis (Kaplan-Meier Curves): A Method to Predict the Future. Arq Bras Oftalmol (2020) 83(2):V–VII. doi: 10.3953/aboj.2020.0036
89. Chakraborty H, Hossain A. R Package to Extract Intracellular Correlation Coefficient With Confidence Interval for Binary Data. Comput Methods Programs Biomed (2018) 155:85–92. doi: 10.1016/j.cmpb.2017.10.023
90. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the Population Abundance of Tissue-Infiltrating Immune and Stromal Cell Populations Using Gene Expression. Genome Biol (2016) 17(1):218. doi: 10.1186/s13059-016-1070-5
91. Chen Y, Chen H, Mao B, Zhou Y, Shi X, Tang L, et al. Transcriptional Characterization of the Tumor Immune Microenvironment and its Prognostic Value for Locally Advanced Lung Adenocarcinoma in a Chinese Population. Cancer Manag Res (2019) 11:9165–73. doi: 10.2147/CMAR.S209571
92. Inglis PW, Pappas MCR, Resende LV, Grattapaglia D. Fast and Inexpensive Protocols for Consistent Extraction of High Quality DNA and RNA From Challenging Plant and Fungal Samples for High-Throughput SNP Genotyping and Sequencing Applications. PloS One (2018) 13(10):e0206885. doi: 10.1371/journal.pone.0206885
93. Garcia-Elias A, Alloza L, Puigdecantet E, Nonell L, Tajes M, Currado J, et al. Defining Quantification Methods and Optimizing Protocols for Microarray Hybridization of Circulating Micrornas. Sci Rep (2017) 7(17):7725. doi: 10.1038/s41598-017-08134-3
94. Shaurova T, Zhang L, Goodrich DW, Hershberger PA. Understanding Lineage Plasticity as a Path to Targeted Therapy Failure in EGFR-Mutant Non-Small Cell Lung Cancer. Front Genet (2020) 11:281. doi: 10.3389/fgene.2020.00281
95. To KKW, Fong W, Cho WCS. Immunotherapy in Treating EGFR-Mutant Lung Cancer: Current Challenges and New Strategies. Front Oncol (2021) 11:635007. doi: 10.3389/fonc.2021.635007
96. Roman M, Baraibar I, Lopez I, Nadal I, Roldo E, Rolfo C, et al. KRAS Oncogene in Non-Small Cell Lung Cancer: Clinical Perspectives on the Treatment of an Old Target. Mol Cancer. (2018) 17(1):33. doi: 10.1186/s12943-018-0789-x
97. Desrichard A, Kuo F, Chowell D, Lee KW, Riaz N, Wong RJ, et al. Tobacco Smoking-Associated Alterations in the Immune Microenvironment of Squamous Cell Carcinomas. J Natl Cancer Inst (2018) 110(12):1386–92. doi: 10.1093/jnci/djy060
98. Lei X, Lei Y, Li JK, Du WX, Li RG, Yang J, et al. Immune Cells Within the Tumor Microenvironment: Biological Functions and Roles in Cancer Immunotherapy. Cancer Lett (2020) 470:126–33. doi: 10.1016/j.canlet.2019.11.009
45. Isomoto K, Haratani K, Hayashi H, Shimizu S, Tomida S, Niwa T, et al. Impact of EGFR-TKI Treatment on the Tumor Immune Microenvironment in EGFR Mutation-Positive Non-Small Cell Lung Cancer. *Clin Cancer Res* (2020) 26(8):2037–46. doi: 10.1158/1078-0432.CCR-19-2027

46. Selimovic D, Wahl RU, Ruiz E, Aslam R, Flanagan TW, Hassan SY, et al. Tumor Necrosis Factor-α Triggers Opposing Signals in Head and Neck Squamous Cell Carcinoma and Induces Apoptosis via Mitochondrial- and Non-Mitochondrial-Dependent Pathways. *Int J Oncol* (2019) 55(6):1324–38. doi: 10.3892/ijo.2019.4900

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Xu, Qin, Yuan, Wei, Sun, Qin, Deng, Zheng and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.