Full text open access online (Since 2001)

PRINT: ISSN 0972-3757 ONLINE: ISSN 2456-6330
Int J Hum Genet, 23(4): 274-289 (2023)
DOI: 10.31901/24566322.2023/23.04.879

Cellular Heterogeneity of Lung Adenocarcinoma is Correlated with EGFR Inhibitor Resistance

Kui Zang, Min Wang, Xingxing Zhu, Bin Yao and Ying Huang*

Department of ICU, The Affiliated Huaian No.1 People's Hospital of Nanjing Medical University, No.1, Huanghe West Road, Huaiyin District, Huai'an City 223300, Jiangsu Province, China

KEYWORDS Drug Resistance. EGFR Inhibitor. Lung Cancer. Single Cell RNA-seq. Heterogeneity

ABSTRACT Although EGFR inhibitors have high curative effect on lung adenocarcinoma, EGFR inhibitor resistance is still a difficult problem that must be overcome at present. By integrating bulk RNA-seq data from EGFR inhibitor-sensitive and resistant lung adenocarcinoma cell lines, the researchers identified key genes associated with EGFR inhibitor resistance. By integrating single-cell RNA-seq data from 12 clinical lung adenocarcinoma samples, 22 subpopulations of lung adenocarcinoma cells were identified. This study not only confirmed that the cellular heterogeneity of lung adenocarcinoma is associated with EGFR inhibitor resistance, but also cell subpopulations with EGFR inhibitor resistance properties, such as cluster6, may be closely related to tumor malignancy, suggesting that the study focusing on EGFR inhibitor resistant subpopulations will help to find suitable strategies to overcome EGFR inhibitor resistance.

INTRODUCTION

The most common lung cancers are non-small cell lung cancers (NSCLC), which account for about 80 percent of all lung malignancies, and the common NSCLC is lung adenocarcinoma, which accounts for about 50 percent of NSCLC. The five-year mortality rate of lung adenocarcinoma is very high, ranging from 51 to 99 percent dependent on different stages (Chalela et al. 2017). Whole-exome sequencing confirms that approximately 14 percent of lung adenocarcinomas carry EGFR mutations (Cancer Genome Atlas Research 2014). EGFR inhibitor has been proved to be very effective for treating lung adenocarcinoma carrying EGFR mutants (Cooper et al. 2022). However, acquired tolerance remains one of the major obstacles to the treatment of lung adenocarcinoma with EGFR inhibitors (Cooper et al. 2022). At present, it is urgent to overcome the obstacle. Comprehensive elucidation of the mechanisms of resistance to EGFR inhibitors is necessary to overcome this obstacle using EGFR inhibitors to treat lung adenocarcinoma. This study plans to explore the relationship between cellular heterogeneity of lung adenocarcinoma and EGFR inhibitor resistance.

Address for correspondence:
Ying Huang,
Department of ICU,
The Affiliated Huaian No.1 People's Hospital of Nanjing Medical University,
No.1, Huanghe West Road, Huaiyin District,
Huai'an City 223300, Jiangsu Province, China
E-mail: huangying5249@163.com

Lung adenocarcinoma has a highly complex heterogeneity, including molecular, cellular, and histopathological heterogeneity (de Sousa and Carvalho 2018). Elucidating heterogeneity is critical for accurate diagnosis and personalized treatment of lung adenocarcinoma. Single-cell RNA-seq offers the opportunity to elucidate the cellular heterogeneity of lung adenocarcinoma. Clinical tumor samples by single-cell RNA sequencing can identify the subpopulations of tumor and different immune cell populations in tumors and the tumor microenvironment, including lung adenocarcinoma (Wu et al. 2021a). To the best of our knowledge, there is short of reports on the relationship between the cellular heterogeneity of lung adenocarcinoma and its resistance to EGFR inhibitors.

Objective

In the present study, the researchers attempted to explore the relationship between cell heterogeneity and EGFR inhibitor resistance in lung adenocarcinoma by integrating bulk RNA-seq and single-cell RNA-seq data. The key genes and gene oncology pathways related to EGFR inhibitor resistance of lung adenocarcinoma were obtained by comparing bulk RNA-seq data from EGFR inhibitor-sensitive and resistant lung adenocarcinoma cell lines. Integrating single-cell RNA-seq data

with these key genes and pathways identified and defined a key lung adenocarcinoma subpopulation associated with EGFR inhibitor resistance. This subpopulation expresses stem cell marker molecules. The genes related to EGFR inhibitor resistance expressed in this subpopulation are closely associated with the survival time of lung adenocarcinoma patients.

MATERIAL AND METHODS

Bulk RNA-seq Data Analysis

Bulk RNA-seq data were downloaded from Gene Expression Omnibus (GEO) (GSE162045, GSE162002 (Rossillo and Ringstad 2020), GSE 199627 (Wu et al. 2022)). After downloading, the data were read into R. The EGFR inhibitor gefitinib resistance and non-resistance data were selected. Gene symbols were fetched from org.Hs.eg.db according to ENSEMBL or entrez ID. The NA values were removed. And then the differential expression (log2 (FC)) values between the EGFR inhibitor resistance and non-resistance were calculated. Genes with log2 (FC) values greater than or equal to 1 were selected as EGFR inhibitor-resistant up-regulated genes, while genes with log2 (FC) values less than or equal to -1 were selected as EGFR inhibitor-resistant down-regulated genes. These genes were used for subsequent further analysis.

Single Cell RNA-seq data Analysis

Single cell RNA-seq data were downloaded from Gene Expression Omnibus (GEO) (GSE189357, GSE203360 (He et al. 2022)). After downloading, the data were input into R, the gene names were obtained from org. Hs.eg.db according to ENSEMBL or entrez id, the NA values were removed, and a two-dimensional matrix of cells and genes was generated. The Seurat (Hao et al. 2021) software package was used to detect and excluded cells with high ratio of mitochondrial genes, and then a series of calculations were performed using SCTransform, RunPCA, FindNeighbors, RunUMAP, and FindClusters functions to obtain cell map. The tumor cells were extracted by using the markers (NASPA and KRT7) of lung adenocarcinoma to form a subset of tumor cell map. Using SelectIntegrationFeatures, PrepSCTIntegration, FindIntegrationAnchors, IntegrateData functions, after a series of calculations, tumor cells were integrated into a Seurat object. Then the clusters were obtained by using SCTransform, RunPCA, FindNeighbors, RunUMAP, FindClusters functions to perform a series of calculations. Further the residual T cells, B cells, macrophages, neutrophils, monocytes, and dendritic cells were excluded according to specific markers of these cells, and then final tumor cell map was obtained by using SCTransform, RunPCA, FindNeighbors, RunUMAP, FindClusters functions to perform a series of calculations. The specific positive or negative markers of tumor cell subpopulations were obtained by using the FindAllMarkers function. The AverageExpression function was used to obtain the average expression of different tumor cell subpopulations.

Gene Oncology Analysis

For bulk RNA-seq data, gene oncology analysis of up- or down-regulated genes was performed in R with the clusterProfiler (Wu et al. 2021b) package using the enrichGO function. The enrichment graphs were made by using the dotplot function. For single-cell RNA-seq data, the genes were sorted in descending order based on differential expression, and then the genes and their expressions were integrated to generate a gene list. Gene oncology analysis was performed on the gene list using the gseGO function. Figures were prepared using the gseaplot2 function. Similar to bulk RNA-seq data analysis, the enrichGO function was used for down-regulated genes of single cell RNA-seq data analysis.

Venn Diagram Analysis

Using the VennDetail package (https://github.com/guokai8/VennDetail), Venn diagram analysis was performed on the obtained target genes in R, and the Venn plots were created with the plot function. All the analysis results were obtained by using the detail function, and the different shared gene information was obtained by using the getSet function.

Survival Analysis

The survival time and related gene expression data of lung adenocarcinoma patients were downloaded from protein atlas (https://www.proteinatlas.

org/). Survival time and associated gene expression data were extracted in R. The optimal cutoff value was calculated using surv_cutpoint function in R. According to the cutoff value of related genes, lung adenocarcinoma patients were divided into two groups with high or low expression of related genes. According to the survival status of lung adenocarcinoma patients, the patients were divided into two groups: death and censused. After normalization of the data, survival analysis was performed in R using the survivor and survminer package. Statistical significance was calculated using the Chisq method, and p_{vale} less than 0.05 was considered statistically significant. The survival curves were prepared using the ggsurvplot function.

RESULTS

Genes and Gene Oncology Pathways Related to EGFR Inhibitor Resistance in Lung Adenocarcinoma

In clinic, EGFR inhibitors are widely used to treat lung adenocarcinomas, but one of the significant drawbacks easily leads to resistance. The current study integrated three EGFR inhibitor (gefitinib)-resistant lung adenocarcinomas and four bulk RNA-seq data to try to identify key EGFR inhibitor resistance-related genes and gene oncology pathways. It was found that 57 genes were 2-fold up-regulated in the three resistant cell lines and four experiments, and more than 300 genes were confirmed to be 2-fold up-regulated in at least three experiments (Fig. 1A). The above genes were used for gene oncology pathway analysis, and found that these genes are involved in cell morphology, cell migration, and extracellular matrix formation, such as collagen-containing extracellular, cell leading raft, basal part of cell, sarcomere, and myofibril (Fig. 1B). At the same time, 50 genes were found to be 2-fold down-regulated in three resistant cell lines and four experiments, and more than 500 genes were confirmed to be 2-fold down-regulated in at least three experiments (Fig. 2A). The above genes were also used for gene oncology pathway analysis, and it was found that the genes were enriched in the gene oncology pathways associated with cell-cell interaction, including cell-cell junction, and tight junction as well as cell membrane-related pathways, including basal plasma membrane (Fig. 2B).

Cellular Heterogeneity of Lung Adenocarcinoma

The cellular heterogeneity of lung adenocarcinoma may be involved in EGFR inhibitor resistance. To address this question, the current study used single-cell RNA-seq data to explore the cellular heterogeneity of lung adenocarcinoma. The single-cell RNA-seq data of 12 lung adenocarcinomas from two independent investigations were integrated to obtain 12,945 lung adenocarcinoma cells, which were divided into 22 subpopulations (Fig. 3A), showing that lung adenocarcinomas have extensive cellular heterogeneity. These lung adenocarcinoma cells all expressed lung adenocarcinoma markers such as NAPSA (Fig. 3B) and KRT7 (Fig. 3C). Each subpopulation has its own markers, and its top 5 marker is sufficient to distinguish these subpopulations (Fig. 4A). To better characterize these cell subpopulations, the single marker that best distinguish these subpopulations were identified, including NAPSA (pan), SLC1A7 (cluster0), RPL10P9 (cluster1), SPINK1 (cluster2), MEG3 (cluster3), CEACAM5 (cluster4), WIF1(cluster5), AGER (cluster6), MSLN (cluster7), CDKN2 A (cluster8), CRABP2 (cluster9), PLAT(cluster10), MTRNR2L1 (cluster11), SRGN(cluster12), RPS4Y1 (cluster13), SFTPD (cluster14), CHI3L1(cluster15), CD69 (cluster16), C3 (cluster17), KRTAP3-1(cluster18), NNMT (cluster19), SOD3(cluster20), IGLC2 (cluster21) (Fig. 4B).

Characterizing Subpopulations of Lung Adenocarcinomas Resistant to EGFR Inhibitors

In order to identify the subpopulations of lung adenocarcinomas that mainly contribute to EGFR inhibitor resistance, the differential expression of the core genes associated with EGFR inhibitor resistance among the cell subpopulations was investigated. Twelve of the 57 key genes up-regulated which were related to EGFR inhibitor resistance were positive markers for subpopulations of lung adenocarcinoma (Fig. 5A), of which at least 6 were mainly expressed in cluster6 (NAPSA+AGER+), including CYBRD1, DOCK8, COL12A1, STX11, RNF144B and FSTL1 (Fig.5B). At the same time, 19 of the 50 key genes down-regulated which were related to EGFR inhibitor resistance were negative markers for subpopulations of lung adenocarcinoma (Fig. 6A), of which at least 6 genes could hardly be detected in cluster6 ($NAPSA^+AGER^+$), includ-

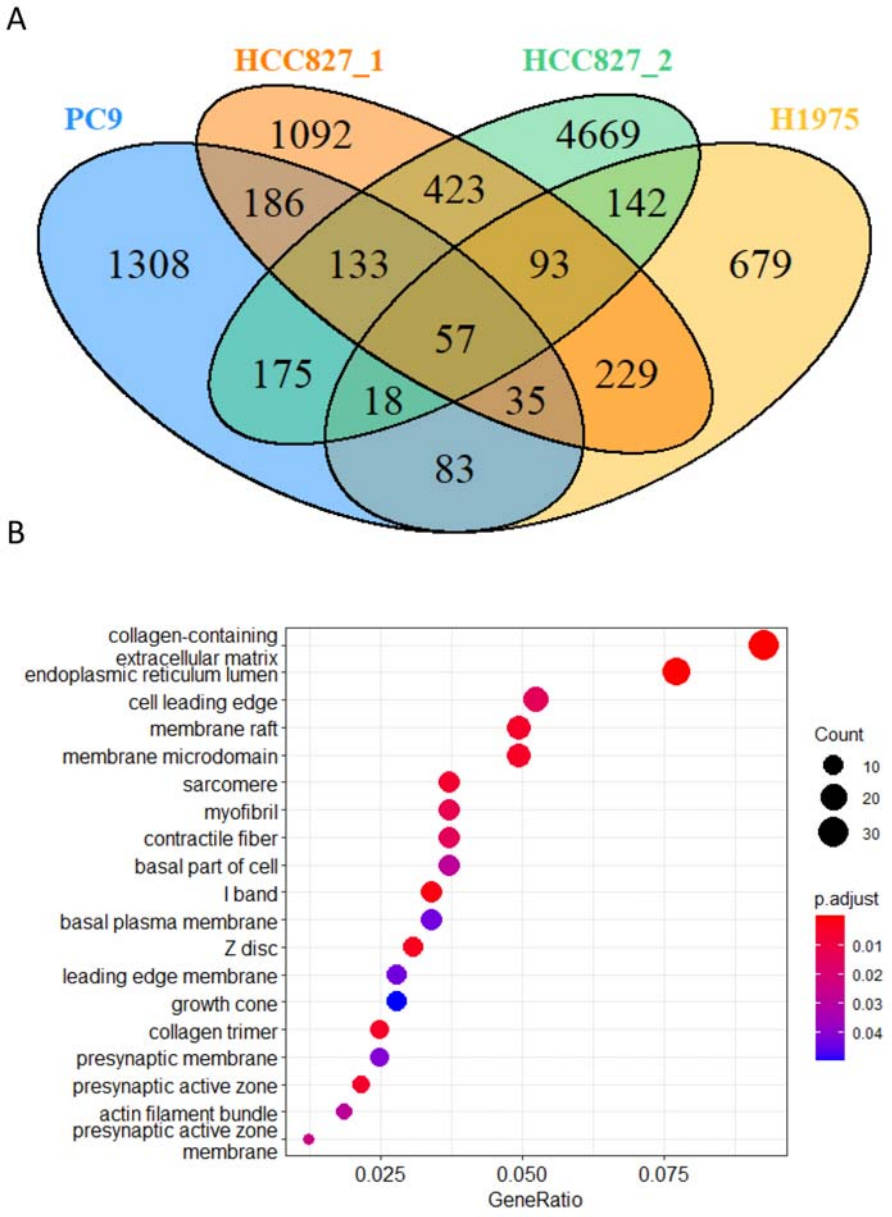

Fig. 1. Up-regulated genes and gene oncology pathways in EGFR inhibitor-resistant lung adenocarcinoma cell lines. Genes with more two-fold up-regulated expression were obtained by comparing batch RNA-seq data of different EGFR inhibitor (gefitinib)-resistant lung adenocarcinoma cell lines. Co-upregulated genes were obtained by Venn analysis (A). The genes which were up-regulated in at least three RNA-seq experiments were used for gene oncology pathway enrichment analysis. The gene oncology pathways with the top 20 enrichment scores are listed (B)

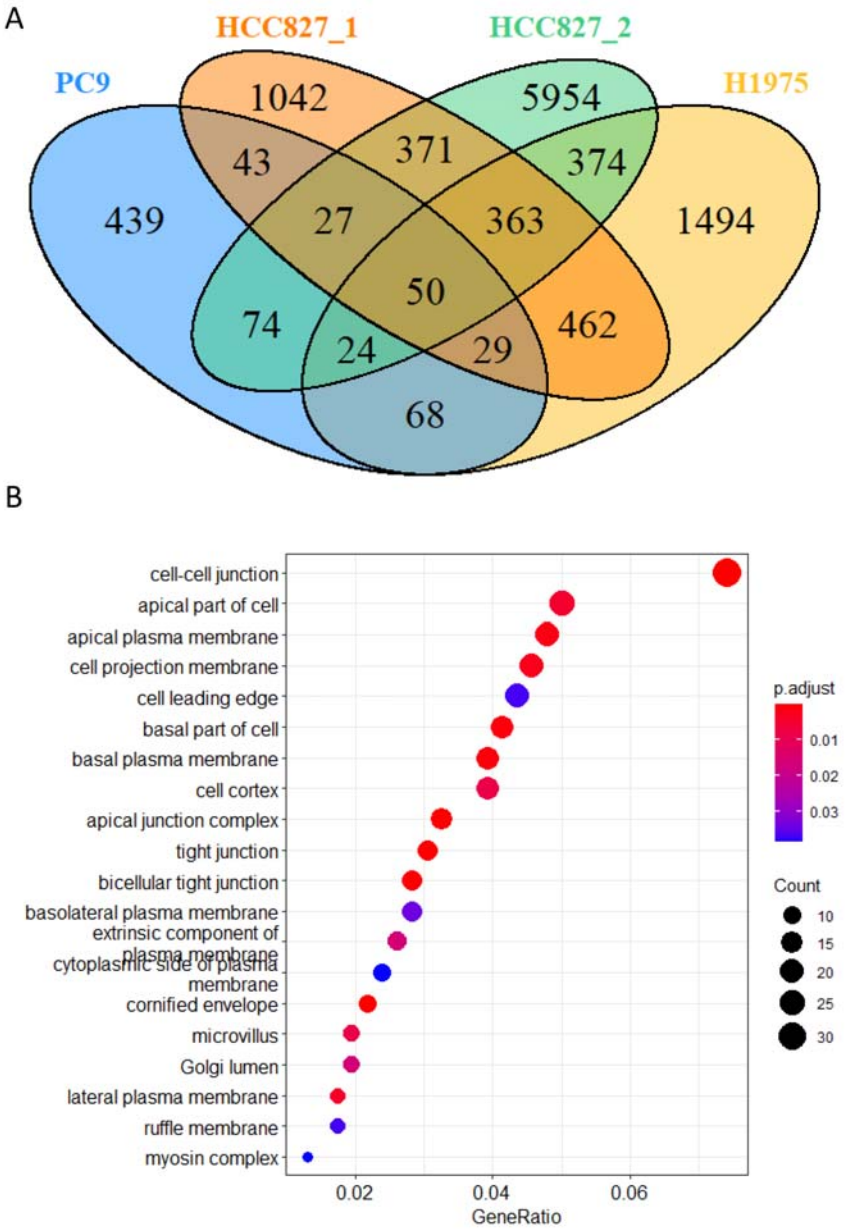

Fig. 2. Down-regulated genes and gene oncology pathways in EGFR inhibitor-resistant lung adenocarcinoma cell lines. Genes with less two-fold down-regulated expression were obtained by comparing batch RNA-seq data of different EGFR inhibitor (gefitinib)-resistant lung adenocarcinoma cell lines. Co-downregulated genes were obtained by Venn analysis (A). The genes which were down-regulated in at least three RNA-seq experiments were used for gene oncology pathway enrichment analysis. The gene oncology pathways with the top 20 enrichment scores are listed (B)

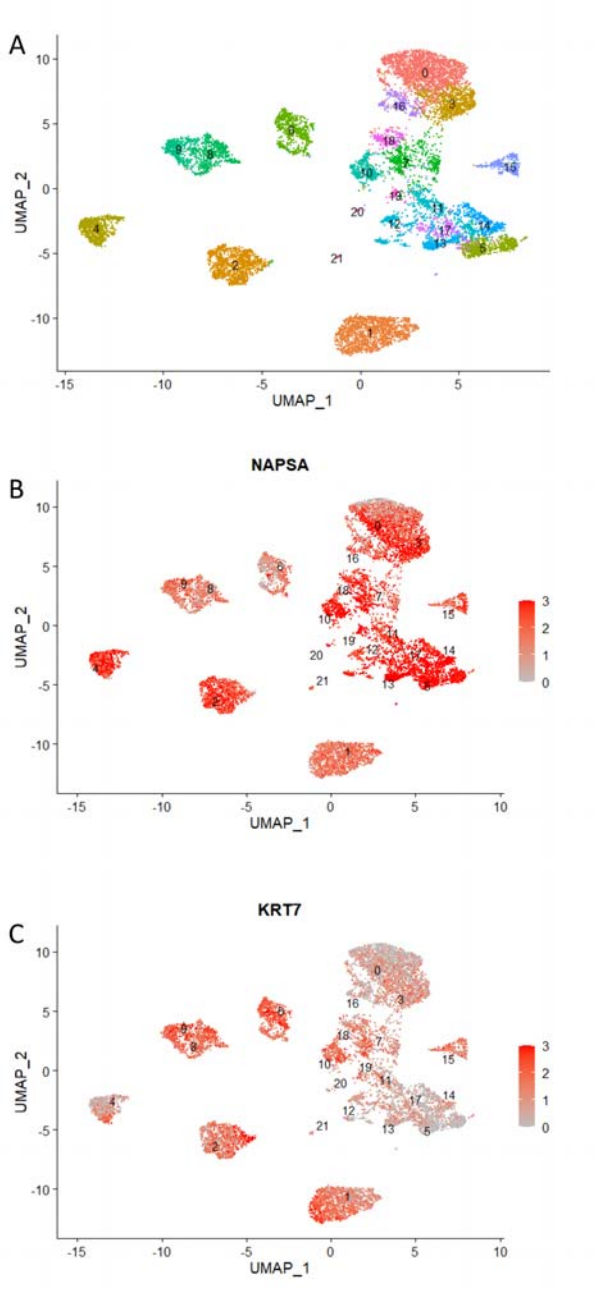

Fig. 3. Single-cell RNA-seq analysis identifies cellular heterogeneity in lung adenocarcinoma. Integrated single-cell RNA-seq data from 12 lung adenocarcinoma samples identified 22 lung adenocarcinoma subpopulations (A). These subpopulations all expressed lung adenocarcinoma marker NAPSA (B) and KRT7 (C)

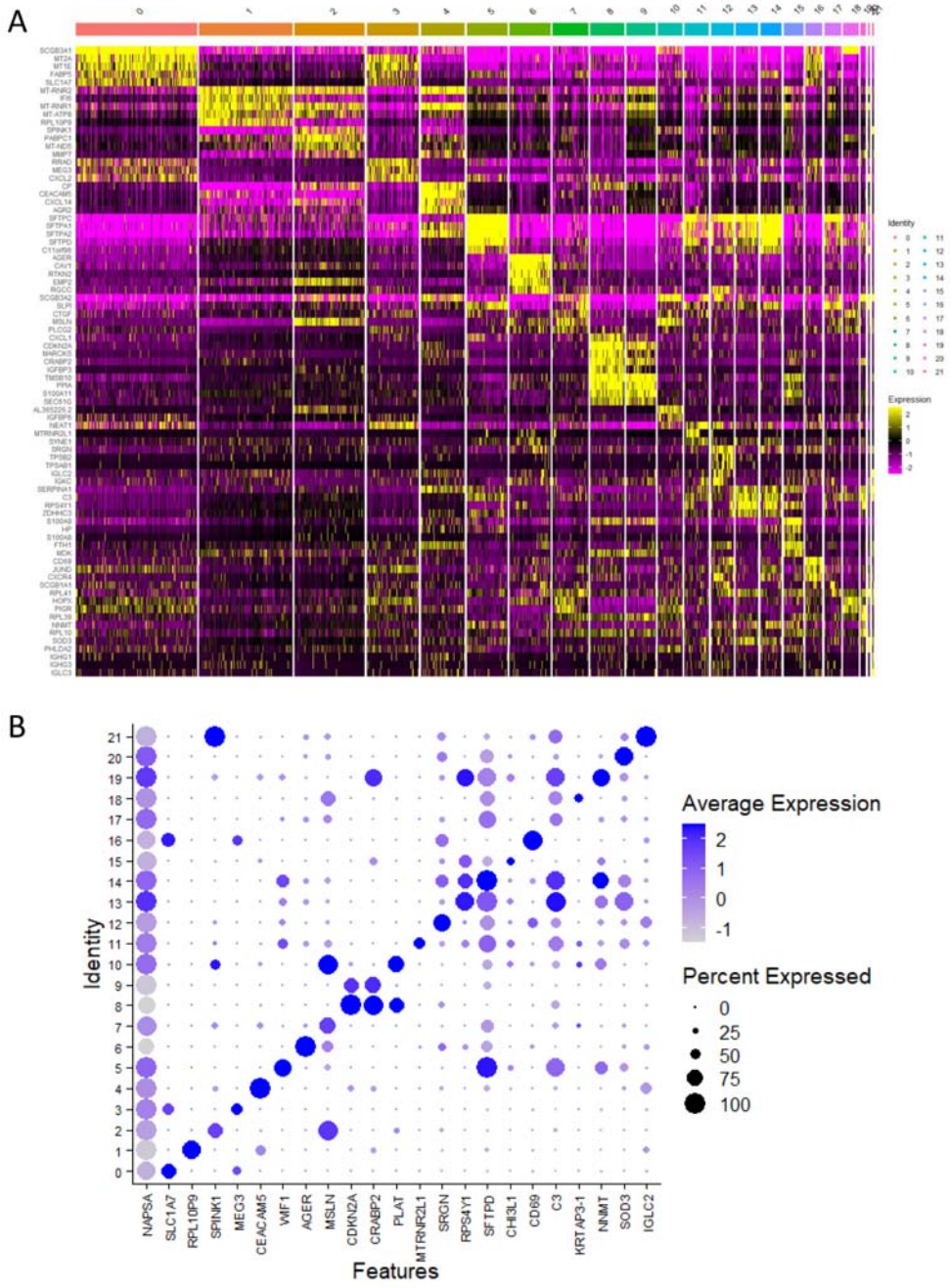

Fig. 4. Characterization of lung adenocarcinoma subpopulations. A. Using the top 5 characteristic genes to make Heatmap, the results showed that top 5 genes are sufficient to distinguish these subpopulations of lung adenocarcinoma. B. The dotplot indicates a single marker that can distinguish these subpopulations of lung adenocarcinoma

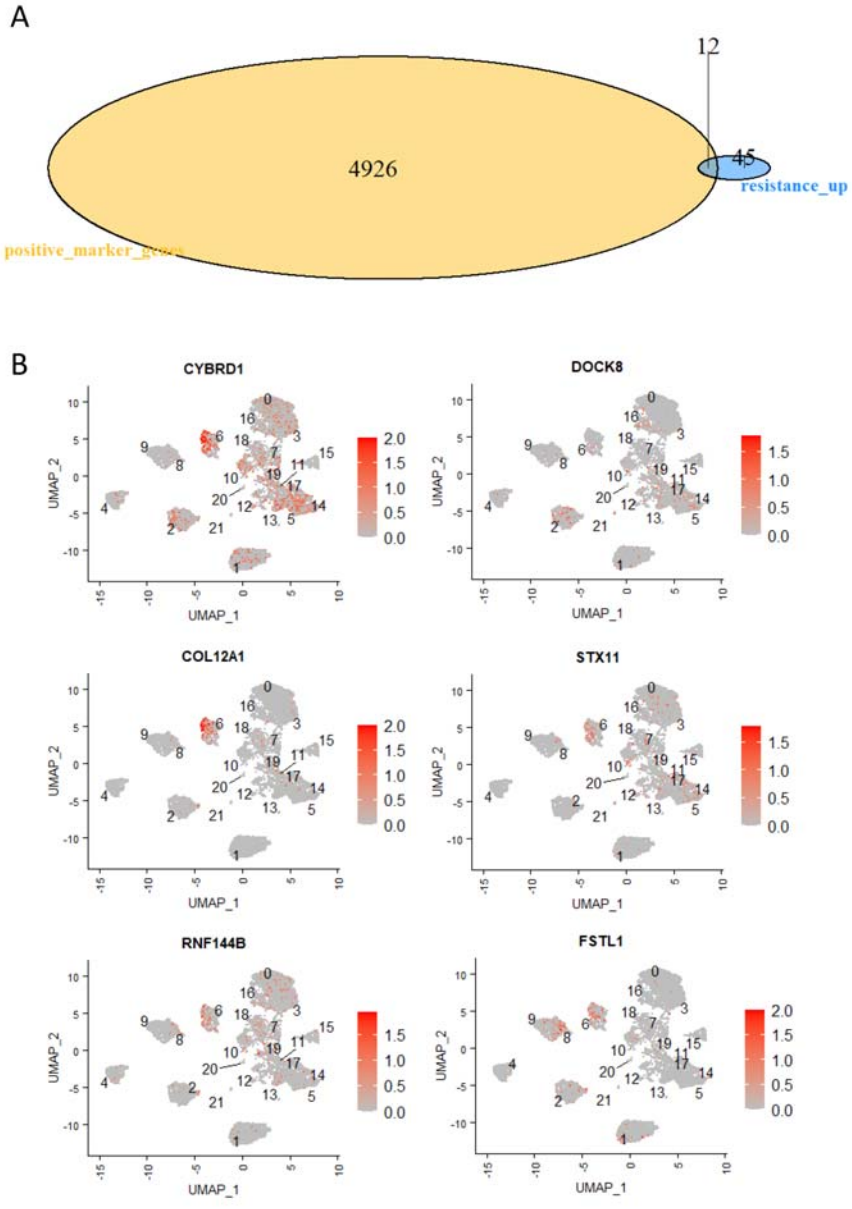

Fig. 5. Up-regulated genes in EGFR inhibitor-resistant lung adenocarcinoma were expressed in lung adenocarcinoma subpopulations. A. Venn analysis revealed that twelve genes are the positive markers of lung adenocarcinoma subpopulations, which were all up-regulated in EGFR inhibitor-resistant lung adenocarcinomas. B. Of the twelve markers that were up-regulated in EGFR inhibitor-resistant lung adenocarcinomas, six were mainly expressed in cluster6, which are NAPAS^{+}AGER^{+} cells

ing ETV1, COL17A1, VIPR1, TMPRSS4, CRABP2 and SPRY4 (Fig. 6B). The above results indicate that cluster6 ($NAPSA^+AGER^+$) is the main subpopulation involved in EGFR inhibitor resistance in lung adenocarcinoma.

In order to further confirm the role of cluster6 in EGFR inhibitor resistance of lung adenocarcinoma, the gene enrichment of cluster6 relative to other subpopulations was analyzed, and it was found that EGFR inhibitor resistance-related pathways were positively enriched in cluster6, $NAPSA^+AGER^+$ cells, including extracellular matrix and cell migration related pathways, such as cell leading edge, collagen-containing extracellular matrix, myofibril, contractile, and sarcomere (Fig. 7A and B). Meanwhile, EGFR inhibitor-resistant down-regulated pathways were also down-regulated in cluster6, $NAPSA^+AGER^+$ cells, especially those associated with cell membranes, including apical part of cell, apical plasma membrane, basal part of cell, and basal plasma membrane (Fig. 7C). These evidences were further confirmed that cluster6 contributes to EGFR inhibitor resistance of lung adenocarcinoma.

Since cluster6, $NAPSA^+AGER^+$ cells, is mainly responsible for EGFR inhibitor resistance of lung adenocarcinoma, the characteristics of cluster6 were further investigated. Comparing the markers of cluster6 with those of stem cells, the researchers found that eight stem cell markers were mainly expressed in cluster6 (Fig. 8A), of which included CDK6, CCND2, GJA1 and ID1 (Fig. 8B). Therefore, these results suggest that cluster6, $NAPSA^+AGER^+$ cells, have some of stem cell properties, that is, cluster6 is likely stem cells of lung adenocarcinoma.

Expression of EGFR Resistance-related Genes in EGFR Inhibitor-resistant Subpopulations of Lung Adenocarcinoma is Associated with Survival Time in Lung Adenocarcinoma Patients

In order to elucidate the importance of revealing the lung adenocarcinoma tolerance-related cell subpopulations in the current study, the researchers investigated the relationship between the expression of EGFR inhibitor tolerance-related genes mainly expressed in cluster6 and the survival time of lung adenocarcinoma patients. It was found that the expression of genes positively correlated with EGFR inhibitor resistance of lung adenocarcinoma was inversely correlated with the survival time of lung adenocarcinoma patients (Fig. 9A). For example, the survival median of lung adenocarcinoma patients with high COL12A1 expression was 3.53 years, while the survival median of lung adenocarcinoma patients with low COL12A1 expression was 4.90 years, and the p_{value} was 0.0096. The survival median of lung adenocarcinoma patients with high FSTL1 expression was 3.44 years, while the survival median of lung adenocarcinoma patients with low FSTL1 expression was 4.11 years, and the p_{value} was 0.23 (Fig. 9A).

Next, the researchers investigated the relationship between the expression of genes negatively correlated with EGFR inhibitor tolerance of lung adenocarcinoma and the survival time of lung adenocarcinoma patients. The results showed that the expression of EGFR inhibitor resistance genes of lung adenocarcinoma, which was barely detectable in cluster6, was positively correlated with the survival time of lung adenocarcinoma patients (Fig. 9B). For example, the survival median of lung adenocarcinoma patients with high expression of ETV1 was 5.95 years, while the survival median of lung adenocarcinoma patients with low ETV1 expression was 3.36 years, and the p_{value} was 0.0013. The survival median of lung adenocarcinoma patients with high expression of SPRY4 was 4.72 years, while the survival median of patients with lung adenocarcinoma with low expression of SPRY4 was 3.33 years, and the p_{value} was 0.021. The survival median of patients with lung adenocarcinoma with high expression of TMPRSS4 was 4.72 years, while the survival median of patients with lung adenocarcinoma with low expression of TMPRSS4 was 3.33 years, and p_{value} was 0.067. The survival median of lung adenocarcinoma patients with high VIPR1 expression was 7.34 years, while the survival median of lung adenocarcinoma patients with low VIPR1 expression was 3.53 years, and p_{value} was 0.0025. The above results indicated that the EGFR inhibitor resistant subpopulation is likely to be associated with malignance of lung adenocarcinoma and of great significance for the prevention and treatment of lung adenocarcinoma.

DISCUSSION

Although EGFR inhibitors are very effective in the treatment of lung adenocarcinoma, tumor resistance must be overcome. At present, tolerance mechanisms mainly include on-target and off-target. The former induces EGFR mutation, making it no longer sensitive to inhibitors, while the latter

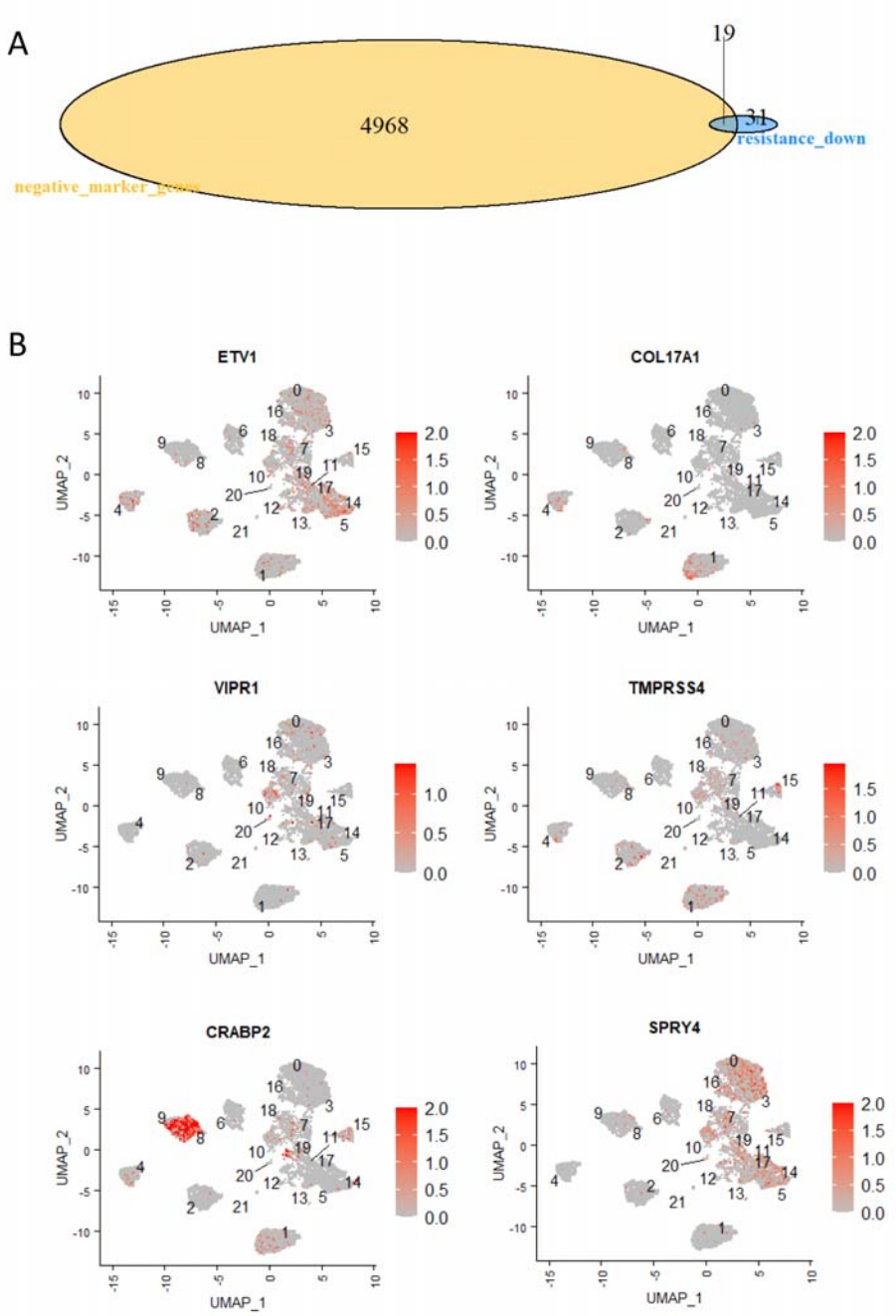

Fig. 6. Down-regulated genes in EGFR inhibitor-resistant lung adenocarcinoma were expressed in lung adenocarcinoma subpopulations. A. Venn analysis revealed that 19 genes are the negative markers of lung adenocarcinoma subpopulations, which were all down-regulated in EGFR inhibitor-resistant lung adenocarcinomas. B. Of the 19 markers that were down-regulated in EGFR inhibitor-resistant lung adenocarcinomas, at least six were negative in cluster6, which are NAPAS$^+$AGER$^+$ cells

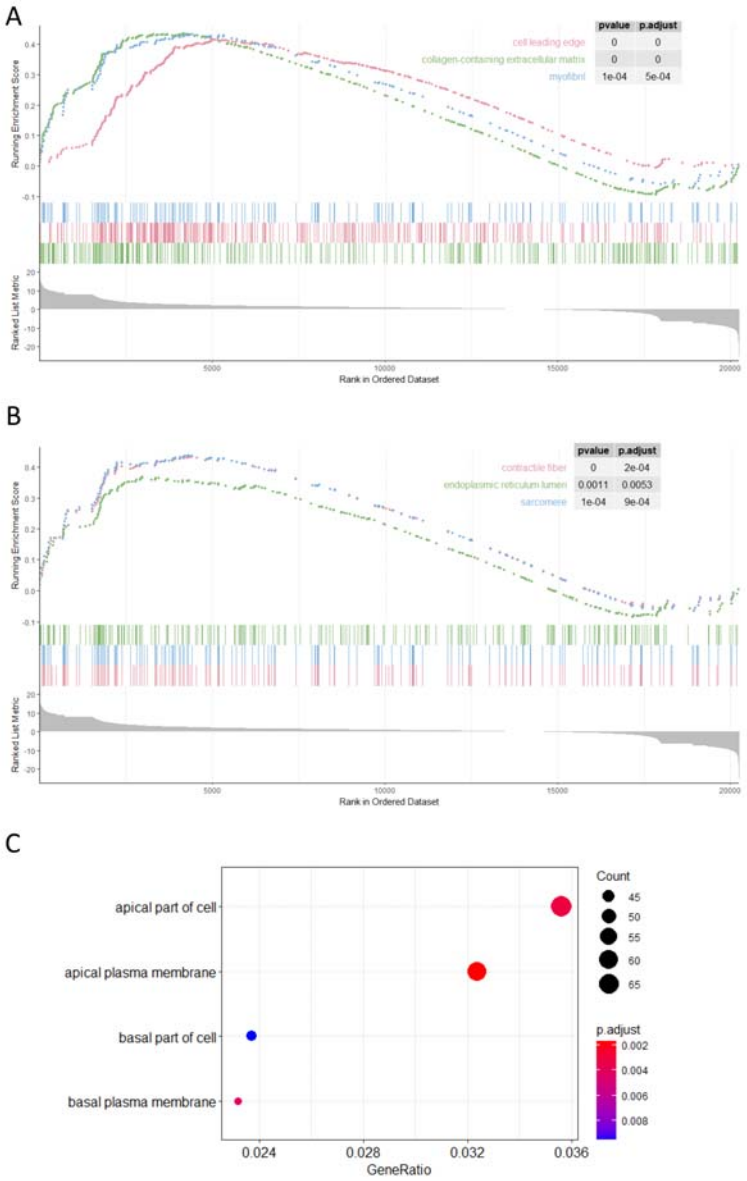

Fig. 7. Gene oncology pathways enriched in EGFR inhibitor-resistant cells were also enriched in lung adenocarcinoma cluster6. Compared with lung adenocarcinoma subpopulation 0, differentially expressed genes in lung adenocarcinoma cluster6 were obtained. Differentially expressed genes were used for gene oncology pathway analysis, and the results revealed that the gene oncology pathways that were positively enriched in EGFR inhibitor-resistant lung adenocarcinoma were also positively enriched in cluster6 (A and B). Compared with lung adenocarcinoma subpopulation 0, genes that were less than twice expressed in lung adenocarcinoma cluster6 were used for gene oncology pathway analysis, and the results revealed that the reduced gene oncology pathways in EGFR inhibitor-resistant lung adenocarcinoma was also decreased in adenocarcinoma cluster6 (C)

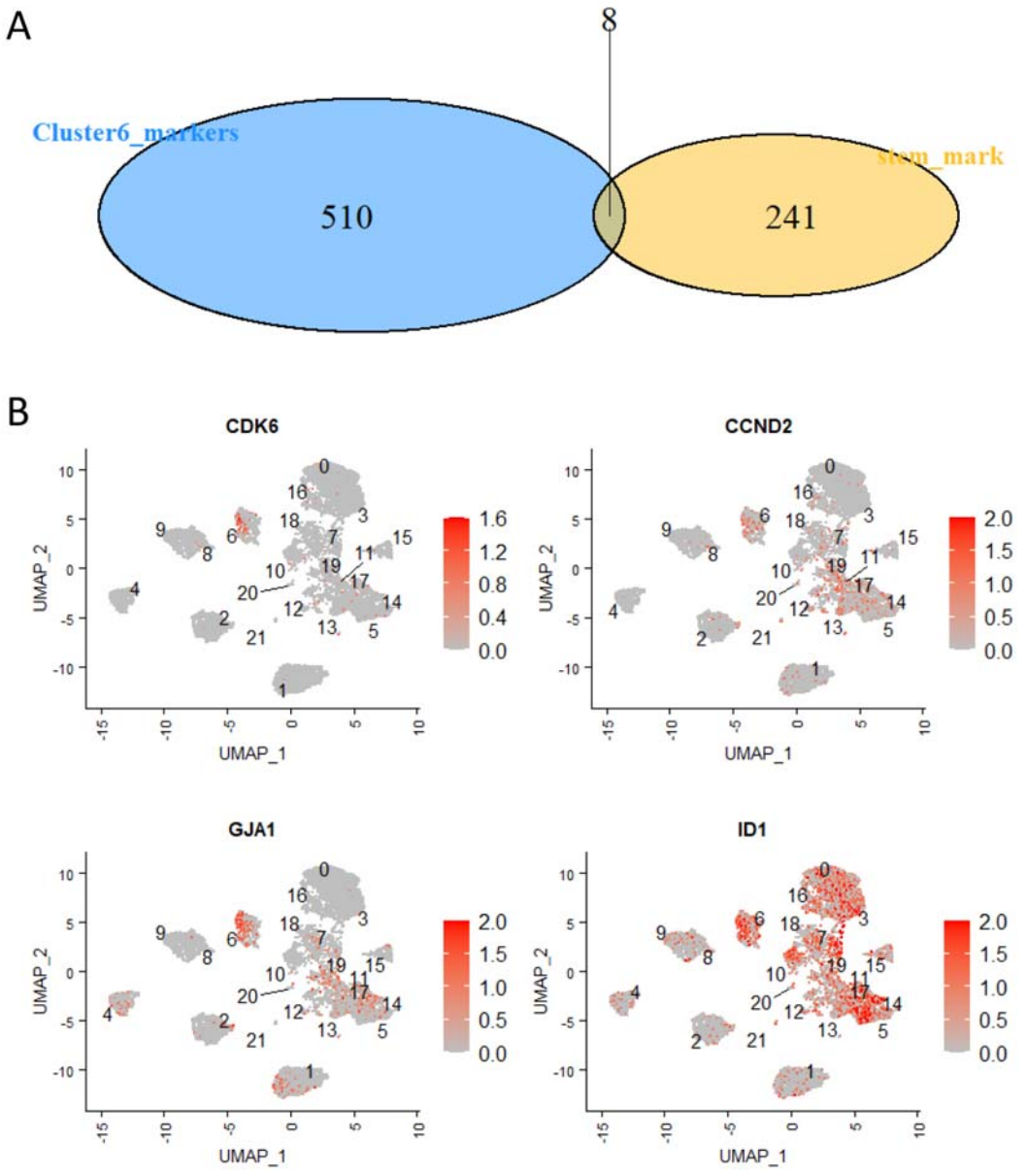

Fig. 8. Lung adenocarcinoma cluster6 with stem cell properties. A. Venn analysis revealed that stem cell markers are also lung adenocarcinoma cluster6 markers. B. Stem cell markers were majorly expressed in lung adenocarcinoma cluster6

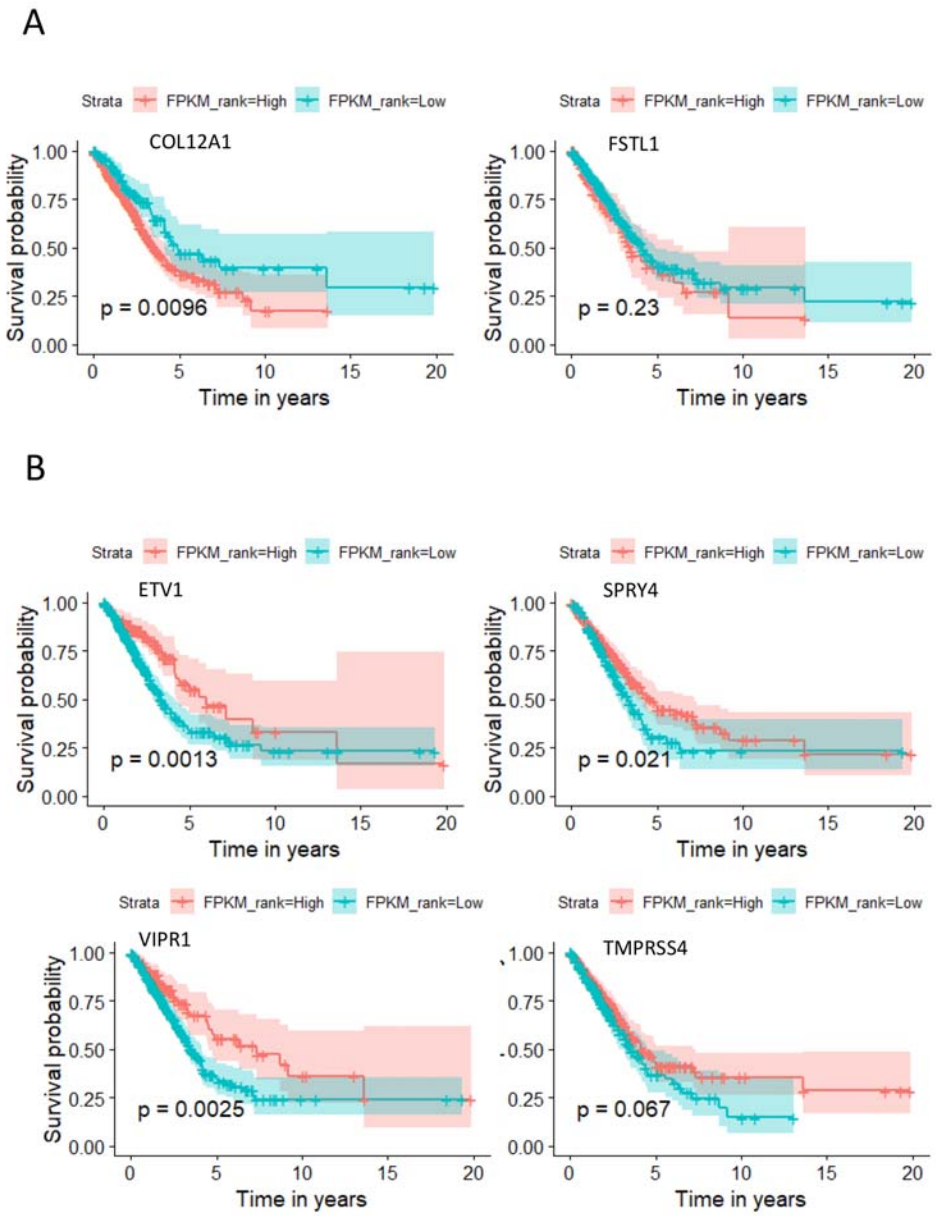

Fig. 9. EGFR inhibitor resistance-related genes expressed in lung adenocarcinoma cluster6 were associated with survival time of lung adenocarcinoma patients. A. The expression of genes positively correlated with EGFR inhibitor resistance in lung adenocarcinoma was inversely correlated with survival time, and these genes were also mainly expressed in lung adenocarcinoma cluster6. B. The expression of genes negatively associated with EGFR inhibitor resistance in lung adenocarcinoma was positively correlated with survival time of lung adenocarcinoma patients, and the expression of these genes was barely detectable in lung adenocarcinoma cluster6

increases the expression of downstream genes of EGFR, thereby skipping the control of EGFR signaling (Cooper et al. 2022). Integrating bulk RNA-seq data from EGFR inhibitor-sensitive and resistant lung adenocarcinoma cells, the researchers found that EGFR inhibitor resistance-related genes are closely associated with cell membrane, cell shape, cell motility, and extracellular matrix. There is no direct evidence for the relationship between the cell membrane and EGFR inhibitor resistance, but the cell membrane can affect the behavior of EGFR such as microscopic EGFR clusters on the membrane and endocytosis after stimulation with ligands (Bag et al. 2015). Consistent with the current findings, EGFR inhibitor-resistant lung adenocarcinoma cells changes to spindle-cell shape, but the effect of the change on EGFR inhibitor tolerance still needs to be further studied (Kim et al. 2013). In the absence of genetic alterations, the extracellular matrix can promote the transformation of EGFR inhibitor-sensitive lung adenocarcinoma cells into EGFR inhibitor-resistant cells, which is dependent on the collagen receptor beta1 integrin (Wang et al. 2019). Beta3 integrin also can enhance EGFR inhibitor resistance of NSCLC (Sun et al. 2022). Current studies have identified COL12A1 as a key gene for EGFR inhibitor resistance of lung adenocarcinoma. Elevated expression of MET and its ligands promotes resistance to EGFR inhibitors, while MET signaling increases cell migration, invasion and motility as well as changes cell morphology (Huang and Fu 2015). At the same time, the juxtamembrane domain mutation of MET can further regulate the cytoskeleton, cell motility and migration in NSCLC (Morgillo et al. 2016). Therefore, the key genes and pathways related to EGFR inhibitor resistance which were obtained by integrating bulk RNA-seq data in this study can be used to evaluate whether tumor cells are resistant to EGFR inhibitors.

Single-cell RNA-seq is currently the best method to study cellular heterogeneity. In fact, several studies have used single-cell RNA-seq to explore the cellular heterogeneity of lung adenocarcinoma (Kim et al. 2020; Ma et al. 2019; Wu et al. 2021a; Xing et al. 2021), but these studies have not focused on tumor cells. Therefore, this study focused on the tumor cells themselves. Given that the number of cells affects the accuracy of clustering, this study used the lung adenocarcinoma cell markers NAPSA and KRT7 to extract tumor cells from 12 single-cell RNA-seq data for integration, and the results identified 22 tumor cell subpopulations, indicating that the lung adenocarcinoma cells have a complex cellular heterogeneity. Although multiple markers can be used to effectively define these subpopulations, for the convenience of application, the best single marker has been identified to distinguish these subpopulations, such as SLC1A7 (cluster0), RPL10P9 (cluster1), MEG3 (cluster3), AGER (cluster6), MTRNR2L1 (cluster11), IGLC2 (cluster21). These markers can be used not only to label and isolate cells in scientific research, but also stain tissues to aid in diagnosis and treatment in clinic.

Based on the identified genes and pathways associated with EGFR inhibitor resistance, cluster6 was confirmed to be characterized by EGFR inhibitor resistance. The genes CYBRD, DOCK8, COL12A1, STX11, RNF144B and FSTL1 positively correlated with EGFR inhibitor resistance were mainly expressed in Cluster6, while the negatively correlated gene ETV1, COL17A1, VIPR1, TMPRSS4, CRABP2 and SPRY4 were hardly expressed in Cluster6. At the same time, gene oncology pathways positively related to EGFR inhibitor tolerance were also positively enriched in cluster6, while gene oncology pathways negatively related to EGFR inhibitor tolerance were also negatively enriched in cluster6. Moreover, cluster6 has stem cell properties. Stem cell markers such as CDK6, CCDN2, GJA1 and ID1 are mainly expressed in cluster6. A study has shown that EGFR inhibitor-resistant cancer cells have stem cell manifestation (Shien et al. 2013). Furthermore, induced stem cell phenotype leads to EGFR inhibitor resistance in lung cancer cells (Liu et al. 2015). In turn, inhibition of stem cell properties of tumor cells overcomes EGFR inhibitor resistance (Si et al. 2019). Therefore, the cellular heterogeneity of lung adenocarcinoma is closely related to EGFR inhibitor resistance.

Cluster6 associated with EGFR inhibitor resistance is likely to be associated with survival time in lung adenocarcinoma patients. Expression of genes positively correlated with EGFR inhibitor resistance which were expressed only in cluster6 is inversely correlated with survival time in lung adenocarcinoma patients, such as COL12A1. COL12A1 has been reported as a tumor prognostic gene and therapeutic target (Jiang et al. 2019). In contrast, the expression of genes negatively correlated with EGFR inhibitor resistance, which was barely detectable in cluster6, positively corre-

lated with survival time in lung adenocarcinoma patients, such as SPRY4 and VIPR1. Overexpression of VIPR1 inhibits the growth, migration and invasion of lung adenocarcinoma cells (Zhao et al. 2019). At the same time, it was confirmed that VIPR1 can also be used as a prognostic gene for lung adenocarcinoma (Jiawei et al. 2020). SPRY4 inhibits the epithelial-mesenchymal transition of NSCLC cells and the proliferation, migration and infiltration of transformed cells (Tennis et al. 2010) while individuals with low SPRY4 expression in colon cancer patients have poor prognosis (Zhou et al. 2016). Therefore, cluster6 is not only related to EGFR inhibitor resistance, but also to tumor malignancy.

CONCLUSION

In conclusion, the researchers integrated bulk RNA-seq data to obtain the key genes and pathways of EGFR inhibitor resistance in lung adenocarcinoma, and integrated single-cell RNA-seq data to obtain the cellular heterogeneity of lung adenocarcinoma. The key genes and pathways of EGFR inhibitor resistance were used to identify cluster6 ($NAPSA^+AGER^+$) with EGFR inhibitor resistance characteristics. Further analysis revealed that Cluster6 also has stem cell properties and may also be associated with tumor malignancy. Therefore, the cellular heterogeneity of lung adenocarcinoma is associated with EGFR inhibitor resistance, and how to overcome this resistance will improve the therapeutic effect of lung adenocarcinoma.

RECOMMENDATIONS

To target the subpopulation of $NAPSA^+AGER^+$ lung adenocarcinoma might be a proper strategy to overcome lung adenocarcinoma resistant to EGFR inhibitors.

ACKNOWLEDGMENTS

The researchers gratefully acknowledge Genomic Data Commons Data Portal and Gene Expression Omnibus.

CONFLICT OF INTERESTS

The authors declare that they have no conflicts of interests.

FUNDING

None.

REFERENCES

Bag N, Huang S, Wohland T 2015. Plasma membrane organization of epidermal growth factor receptor in resting and ligand-bound states. *Biophys J*, 109: 1925-1936.

Cancer Genome Atlas Research N 2014. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*, 511: 543-550.

Chalela R, Curull V, Enriquez C, Pijuan L, Bellosillo B, Gea J 2017. Lung adenocarcinoma: from molecular basis to genome-guided therapy and immunotherapy. *J Thorac Dis*, 9: 2142-2158.

Cooper AJ, Sequist LV, Lin JJ 2022. Third-generation EGFR and ALK inhibitors: Mechanisms of resistance and management. *Nat Rev Clin Oncol*, 19: 499-514.

de Sousa VML, Carvalho L 2018. Heterogeneity in lung cancer. *Pathobiology*, 85: 96-107.

Hao Y, Hao S, Andersen-Nissen E et al. 2021. Integrated analysis of multimodal single-cell data. *Cell*, 184: 3573-3587 e3529.

He Y, Yu F, Tian Y et al. 2022. Single-cell RNA sequencing unravels distinct tumor microenvironment of different components of lung adenocarcinoma featured as mixed ground-glass opacity. *Front Immunol*, 13: 903513.

Huang L, Fu L 2015. Mechanisms of resistance to EGFR tyrosine kinase inhibitors. *Acta Pharm Sin B*, 5: 390-401.

Jiang X, Wu M, Xu X et al. 2019. COL12A1, a novel potential prognostic factor and therapeutic target in gastric cancer. *Mol Med Rep*, 20: 3103-3112.

Jiawei Z, Min M, Yingru X et al. 2020. Identification of key genes in lung adenocarcinoma and establishment of prognostic mode. *Front Mol Biosci*, 7: 561456.

Kim HR, Kim WS, Choi YJ, Choi CM, Rho JK, Lee JC 2013. Epithelial-mesenchymal transition leads to crizotinib resistance in H2228 lung cancer cells with EML4-ALK translocation. *Mol Oncol*, 7: 1093-1102.

Kim N, Kim HK, Lee K et al. 2020. Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat Commun*, 11: 2285.

Liu YN, Chang TH, Tsai MF et al. 2015. IL-8 confers resistance to EGFR inhibitors by inducing stem cell properties in lung cancer. *Oncotarget*, 6: 10415-10431.

Ma KY, Schonnesen AA, Brock A et al. 2019. Single-cell RNA sequencing of lung adenocarcinoma reveals heterogeneity of immune response-related genes. *JCI Insight*, 4.

Morgillo F, Della Corte CM, Fasano M, Ciardiello F 2016. Mechanisms of resistance to EGFR-targeted drugs: lung cancer. *ESMO Open*, 1: e000060.

Rossillo M, Ringstad N 2020. Development of specialized sensory neurons engages a nuclear receptor required for functional plasticity. *Genes Dev*, 34: 1666-1679.

Shien K, Toyooka S, Yamamoto H et al. 2013. Acquired resistance to EGFR inhibitors is associated with a manifestation of stem cell-like properties in cancer cells. *Cancer Res*, 73: 3051-3061.

Si J, Ma Y, Bi JW et al. 2019. Shisa3 brakes resistance to EGFR-TKIs in lung adenocarcinoma by suppressing cancer stem cell properties. *J Exp Clin Cancer Res*, 38: 481.
Sun Q, Lu Z, Zhang Y et al. 2022. Integrin beta3 promotes resistance to EGFR-TKI in non-small-cell lung cancer by upregulating AXL through the YAP pathway. *Cells*, 11.
Tennis MA, Van Scoyk MM, Freeman SV, Vandervest KM, Nemenoff RA, Winn RA 2010. Sprouty-4 inhibits transformed cell growth, migration and invasion, and epithelial-mesenchymal transition, and is regulated by Wnt7A through PPARgamma in non-small cell lung cancer. *Mol Cancer Res*, 8: 833-843.
Wang Y, Zhang T, Guo L, Ren T, Yang Y 2019. Stromal extracellular matrix is a microenvironmental cue promoting resistance to EGFR tyrosine kinase inhibitors in lung cancer cells. *Int J Biochem Cell Biol*, 106: 96-106.
Wu F, Fan J, He Y et al. 2021a. Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. *Nat Commun*, 12: 2540.
Wu L, Yu Y, Xu L, Wang X, Zhou J, Wang Y 2022. TROY modulates cancer stem-like cell properties and gefitinib resistance through emt signaling in non-small cell lung cancer. *Front Genet*, 13: 881875.
Wu T, Hu E, Xu S et al. 2021b. ClusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)*, 2: 100141.
Xing X, Yang F, Huang Q et al. 2021. Decoding the multicellular ecosystem of lung adenocarcinoma manifested as pulmonary subsolid nodules by single-cell RNA sequencing. *Sci Adv*, 7.
Zhao L, Yu Z, Zhao B 2019. Mechanism of VIPR1 gene regulating human lung adenocarcinoma H1299 cells. *Med Oncol*, 36: 91.
Zhou X, Xie S, Yuan C et al. 2016. Lower expression of SPRY4 predicts a poor prognosis and regulates cell proliferation in colorectal cancer. *Cell Physiol Biochem*, 40: 1433-1442.

Paper received for publication in March, 2023
Paper accepted for publication in December, 2023