



## Original article

## Survival analysis in Caucasian pulmonary adenocarcinoma patients based on differential targets between Caucasian and Asian population

Zheng Zhu<sup>a,1</sup>, Zhigang Liang<sup>b,1</sup>, Jichun Tong<sup>c</sup>, Xiaoliang Mao<sup>c</sup>, Yajun Yin<sup>c</sup>, Lydia C. Manor<sup>d,\*</sup>, Zhenya Shen<sup>a,\*</sup><sup>a</sup> Institute for Cardiovascular Science & Department of Cardiovascular Surgery of the First Affiliated Hospital, Soochow University, China<sup>b</sup> Department of Thoracic Surgery, Ningbo First Hospital, Ningbo, Zhejiang 315000, China<sup>c</sup> Department of Cardiovascular Thoracic Surgery, Changzhou No.2 People's Hospital, Nanjing Medical University, Changzhou, China<sup>d</sup> Department of Biology Products, American Informatics LLC, Rockville 20850, USA

## ARTICLE INFO

## Article history:

Received 14 May 2018

Revised 15 May 2018

Accepted 21 May 2018

Available online 22 May 2018

## Keywords:

Pulmonary adenocarcinoma

Caucasian

EGFR

TP53

KRAS

STK11

## ABSTRACT

Ethnicity differences may contribute to the variety of overall survival in pulmonary adenocarcinoma, while the influence of ethnicity relevant somatic driver mutations (ERSDM) profile on Caucasian survival is not well investigated. In this study, we studied epidermal growth factor receptor (EGFR), tumor protein p53 (TP53), Kirsten rat sarcoma 2 viral oncogene homolog (KRAS), and Serine/Threonine Kinase 11 (STK11) to construct the ERSDM profile. Those genes were selected as harboring somatic driver mutations with >10% prevalence and with different occurrence between Caucasian and Asian ethnicity. Clinical information and transcriptome sequencing of 173 Caucasian pulmonary adenocarcinoma patients with matched mutation data are retrieved from TCGA, Kaplan-Meier analyses and Cox proportional-hazards regression models are further used to analyze the effect of the ERSDM profile on overall survival. There is no significant correlation between single gene mutation and overall survival, while patients with less than two mutated genes have a better overall survival compared with those with at least two mutated genes ( $p = 0.034$ ). All of these indicate that multiple mutations in the ERSDM profile may be a negative prognostic factor for overall survival in Caucasian pulmonary adenocarcinoma patients.

© 2018 Production and hosting by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

As the leading cause of cancer death, pulmonary adenocarcinoma is currently the most common type of lung cancer in women, Asians, and people under the age of 45, which accounts for approximately 40% of lung cancers around the globe (Denisenko et al., 2018; Ma et al., 2018; Network, 2014). Because the early symptoms of pulmonary adenocarcinoma are easily omitted and some patients at the first diagnosis will present with distant metastases, the average five-year survival rate is only around 18 percent (Attarian et al., 2017; Dolly et al., 2017).

\* Corresponding authors.

E-mail addresses: [l.manor@gousinfo.com](mailto:l.manor@gousinfo.com) (L.C. Manor), [uuzhenya@aliyun.com](mailto:uuzhenya@aliyun.com) (Z. Shen).<sup>1</sup> These authors contributed equally to this work.

Peer review under responsibility of King Saud University.

In addition to exogenous exposures, inherited genetic alterations, accumulation of somatic genetic events, ethnic differences evidenced by candidate polymorphisms or genome-wide association studies may also contribute to the development and clinical outcome of pulmonary adenocarcinoma (Cai et al., 2017; Schabath et al., 2016a; Zhang et al., 2017). Both epidemiologic research and clinical trials have confirmed that Caucasian ethnicity is not a favorable prognostic factor for overall survival when compared with Asian ethnicity (Soo et al., 2012). Previous comprehensive genomic profiling study also shows that epidermal growth factor receptor (EGFR), tumor protein p53 (TP53), Kirsten rat sarcoma 2 viral oncogene homolog (KRAS), and Serine/Threonine Kinase 11 (STK11) are among the most differential detected somatic driver mutations between Asian and Caucasian ethnicity (Farjah et al., 2009; Liu et al., 2017).

However, the effects of ethnicity relevant somatic driver mutations (ERSDM) profile on the final clinical outcome, such as overall survival, have been investigated by only limited studies. This study aims to explore the ERSDM profile on the survival of Caucasian pulmonary adenocarcinoma.



Production and hosting by Elsevier

## 2. Methods & materials

### 2.1. Data collection

Clinical information and somatic nonsynonymous mutations data from 393 Caucasian pulmonary adenocarcinoma patients were downloaded from the Cancer Genome Atlas (TCGA) data portal (<http://tcgadata.nci.nih.gov>) using cBioPortal (<http://www.cbioportal.org/>). Of them, matched somatic mutation data could be recalled for 173 Caucasian patients, which was included in the downstream analysis. The detailed baseline characteristics can be seen in the [supplemental table 1](#).

### 2.2. Ethnicity relevant somatic driver mutations profile construction

Based on previous comprehensive genomic profiling analysis (Liu et al., 2017), significantly different genes in the prevalence of somatic driver mutations between Caucasian and East Asian patients were selected and further screened with mutation prevalence >10% in TCGA database. Such multiple mutations genes were defined as ethnicity relevant somatic driver mutations profile.

### 2.3. Data analysis

Transcriptome sequencing data for those patients were also downloaded and the estimated expression level was quantified as transcripts per million reads (TPM). Log transformation with base 2 was applied to expression data to meet the normal distribution assumption for downstream statistical tests, and a pseudo number of 0.5 was added to each value to avoid zero value in log transformation. Kaplan-Meier analyses and Cox proportional hazards regression models were used for overall survival (OS) analysis. Hazard ratios for OS were adjusted for baseline characteristics. Unless specified, all analyses were conducted using R, and multiple tests were adjusted using Benjamini-Hochberg approach.

## 3. Results

### 3.1. ERSDM distribution

EGFR, TP53, KRAS, and STK11 were selected as ERSDM profile. Of 173 patients, we observed 39 (23%) without somatic mutation in any of those four genes, 82 (47%) with somatic mutations in one gene, 51 (29%) have two mutated genes, while only one patient harbored mutations in all of four genes. In general, approximately 77% of Caucasian patients with pulmonary adenocarcinoma have at least one somatic mutation on those genes, consistent with their importance in tumorigenesis (Fig. 1).

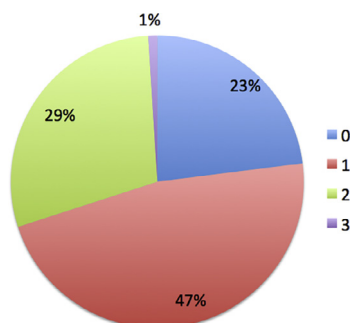


Fig. 1. Distribution of ethnicity relevant somatic driver mutations profile across patients.

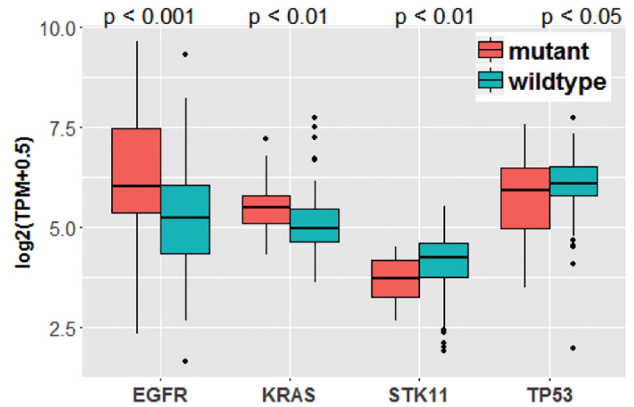


Fig. 2. The expression level of patients with ethnicity relevant somatic driver mutations genes and those without mutations.

### 3.2. Expression pattern

A question of interest was whether somatic mutations could correlate with changes in transcription. To address this, the expression levels of patients with or without mutations in those targeted genes were compared. Interestingly, a significant difference (after multiple test correction) was observed for all four genes, and EGFR and KRAS showed elevated expression when somatic driver mutations were presented. In contrast, tumor suppressor genes such as STK11 and TP53 showed the opposite expression trend (Fig. 2). All of these indicated that mutations of ERSDM profile did alter the relevant transcription.

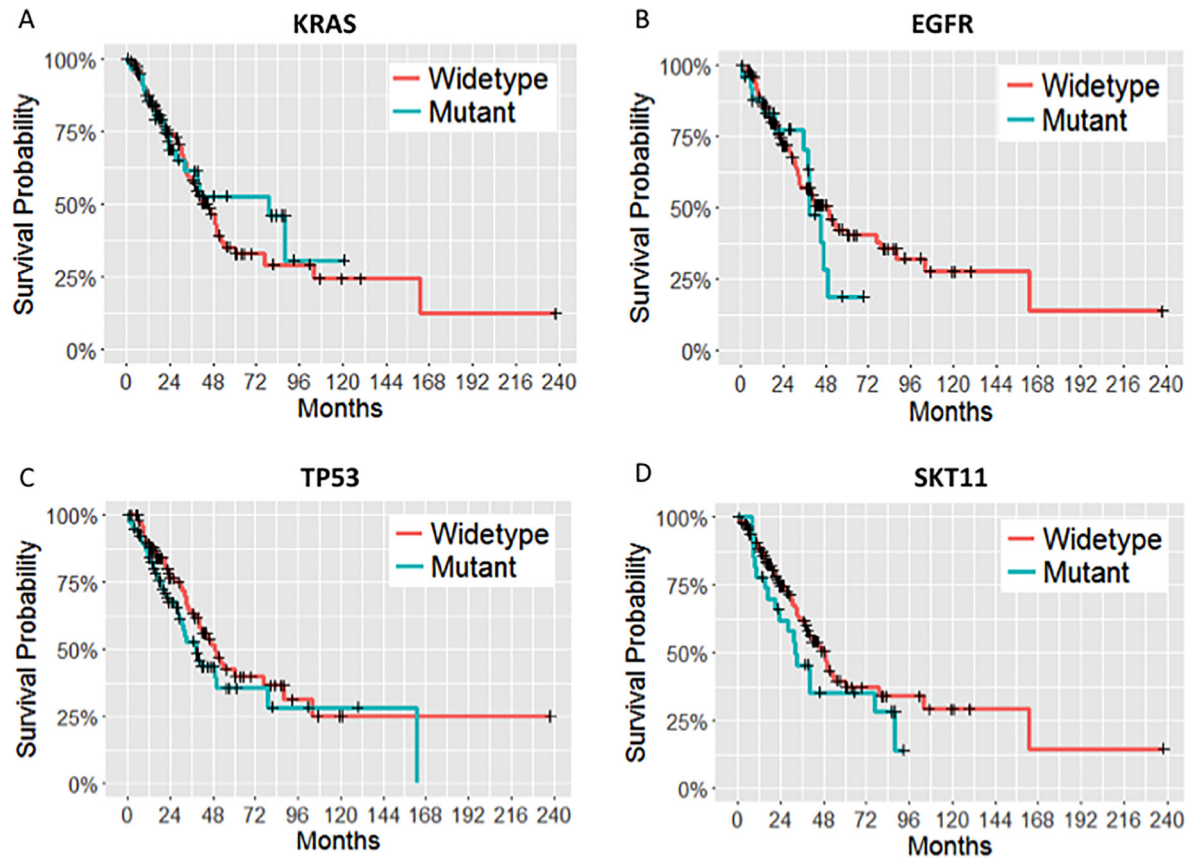
### 3.3. Kaplan-Meier and Cox proportional-hazards regression analysis

Next, we sought to examine the potential prognostic effect of ERSDM profile. To achieve that, Kaplan-Meier analyses were performed. No statistically significant correlation between single genes and overall survival was observed (Fig. 3). Finally, we investigated whether the accumulation of multiple mutations might be a better prognostic biomarker. Indeed, we found patients with less than two mutated genes have a better overall survival compared with those with at least two mutated genes ( $p = 0.034$ ) (Fig. 4). The ethnicity relevant mutations burden would be detrimental to the overall survive.

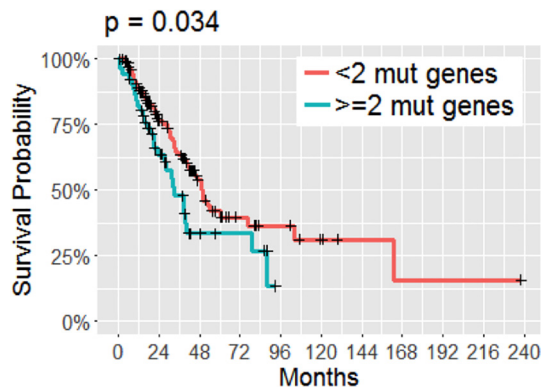
## 4. Discussion

In this study the ERSDM profile (EGFR, TP53, STK11, and KRAS) is constructed as biomarkers to predict survival in Caucasian pulmonary adenocarcinoma populations. Our results demonstrate that no significant correlation is found between single gene mutation and overall survival, while the ethnicity relevant mutations burden and survival rate may show a significant correlation, which indicates that the ERSDM profile can be used in the clinical practice to predict the overall survival, although the precise mechanism still needs further investigated.

It needs to be pointed out that Asian ethnicity and adenocarcinoma histology continues to predict response to Gefitinib (EGFR inhibitors) in patients treated for advanced non-small cell carcinoma of the lung in North America (Araujo and Carbone, 2017; Ho et al., 2005; Schabath et al., 2016b), which indicates that ethnicity does have some relation with pulmonary adenocarcinoma overall survival. What makes it more complex is that modifiable exposures, clinical and behavioral factors, and potential sources of racial and ethnic disparities are conducive to the differences in incidence and mortality in pulmonary adenocarcinoma which are



**Fig. 3.** Correlation of ethnicity relevant somatic driver mutations profile with overall survival. Survival probability comparison among KRAS (A), EGFR (B), TP53 (C), and STK11 (D) mutant phenotypes and the relevant wide type phenotype.



**Fig. 4.** Statistically significant correlation was observed between patients with  $\geq 2$  mutated genes and patients with  $< 2$  genes.

attributed by ethnicity differences (Esnaola et al., 2008; Relin Yang et al., 2010; Ryn, 2002). Ethnic differences associated pulmonary adenocarcinoma incidence and mortality will be altered or reduced if such factors are stratified. It is reported that when adjusted for age, smoking, and other factors, women of different ethnic groups will show no differences in lung cancer incidence or mortality (Home, 2010; Houston et al., 2018; Patel et al., 2016). All of these indicate that when ethnicity difference is considered, more detailed data collection, data cleaning, and data stratify are vital to give an appropriate conclusion and to reduce ambiguities.

Pulmonary adenocarcinoma shows high rates of genomic rearrangement and somatic mutation (FACS et al., 2014). Due to dramatically improved molecular targeted therapies, Gefitinib

therapy can be applied to patients whose tumors harbor somatic activated EGFR mutation. However, most pulmonary adenocarcinomas either lack an identifiable somatic driver mutation, or harbor mutations like KRAS, TP53, and STK11, which are not currently clinically actionable and can only be treated with conventional chemotherapy (Quadrelli and Solís, 2018; Yu and Davidson, 2018). It is worth noting that the prevalence of EGFR and KRAS mutations is relatively low, while tumor suppressor gene STK11 and TP53 abnormalities are relatively high in Caucasian ethnicity when compared with Asian ethnicity (Daga et al., 2015; Liu et al., 2017; Román et al., 2018). The oncogenes and anti-oncogenes mutations difference may result in a paradigm shift, which may contribute to the different clinical outcomes.

It is worth noting that it is really hard to draw a clear distinction between a person's ethnicity might be a pulmonary adenocarcinoma risk factor or a genotypic profile that can determine the drug response, for ERSDM profile does exist. Clinical practice can either be patient-tailored or ethnicity-based. Further investigation about ERSDM profile will give a more clear clue to solve such puzzle.

## 5. Conclusion

Multiple mutations in the ERSDM profile can be used to predict an increased risk of poor survival in Caucasian pulmonary adenocarcinoma patients.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.sjbs.2018.05.023>.

## References

- Araujo, L.H., Carbone, D.P., 2017. Non-small cell lung cancer genomics around the globe: focus on ethnicity. *J. Thoracic Disease* 9, E392.
- Attarian, S., Rahman, N., Halmos, B., 2017. Emerging uses of biomarkers in lung cancer management: molecular mechanisms of resistance. *Ann. Transl. Med.* 5, 377.
- Cai, G., Xiao, F., Cheng, C., Li, Y., Amos, C.I., Whitfield, M.L., 2017. Population effect model identifies gene expression predictors of survival outcomes in lung adenocarcinoma for both Caucasian and Asian patients. *Plos One* 12, e0175850.
- Daga, A., Ansari, A., Patel, S., Mirza, S., Rawal, R., Umrana, V., 2015. Current drugs and drug targets in non-small cell lung cancer: limitations and opportunities. *Asian Pac. J. Cancer Prev.* 16, 4147–4156.
- Denisenko, T.V., Budkevich, I.N., Zhivotovsky, B., 2018. Cell death-based treatment of lung adenocarcinoma. *Cell Death Dis.* 9.
- Dolly, S.O., Collins, D.C., Sundar, R., Papat, S., Yap, T.A., 2017. Advances in the development of molecularly targeted agents in non-small-cell lung cancer. *Drugs* 77, 1–15.
- Esnaola, N.F., Gebregziabher, M., Knott, K., Finney, C., Silvestri, G.A., Reed, C.E., Ford, M.E., 2008. Underuse of surgical resection for localized, non-small cell lung cancer among Whites and African Americans in South Carolina. *Ann. Thoracic Surgery* 86, 220–227.
- FACS, J.A.R.M., DMSc, W.K.H.M., Tsao, A.S., PhD, J.Y.C.M.M., FACS, S.H.B.M.M., 2014. *Somatic Genome Alterations in Human Lung Cancers*. John Wiley & Sons, Inc.
- Farjah, F., Wood, D.E., Rd, Y.N., Vaughan, T.L., Symons, R.G., Krishnadasan, B., Flum, D.R., 2009. Racial disparities among patients with lung cancer who were recommended operative therapy. *Arch. Surg.* 144, 14–18.
- Ho, C., Murray, N., Laskin, J., Melosky, B., Anderson, H., Bebb, G., 2005. Asian ethnicity and adenocarcinoma histology continues to predict response to gefitinib in patients treated for advanced non-small cell carcinoma of the lung in North America. *Lung Cancer* 49, 225.
- Home, C., 2010. Racial/Ethnic disparities and geographic differences in lung cancer incidence – States and the District of Columbia, 1998–2006. *Mmwr. Morb. Mortal Wkly Rep.* 59, 1434–1438.
- Houston, K.A., Mitchell, K.A., King, J., White, A., Ryan, B.M., 2018. Histologic lung cancer incidence rates and trends vary by race/ethnicity and residential county. *J. Thoracic Oncol.* 13.
- Liu, L., Liu, J., Shao, D., Deng, Q., Tang, H., Liu, Z., Chen, X., Guo, F., Lin, Y., Mao, M., 2017. Comprehensive genomic profiling of lung cancer using a validated panel to explore therapeutic targets in East Asian patients. *Cancer Sci.* 108, 2487.
- Ma, J.A., Jiang, S., Hu, C., Xie, Y., Hou, T., 2018. Adjuvant therapy for resected EGFR -mutant non-small-cell lung cancer. *Lancet Oncol.* 19, e124.
- Network, C.G.A., 2014. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 511, 543–550.
- Patel, M.I., Wang, A., Kappahn, K., Desai, M., Chlebowski, R.T., Simon, M.S., Bird, C. E., Corbiesmith, G., Gomez, S.L., Adamscampbell, L.L., 2016. Racial and ethnic variations in lung cancer incidence and mortality: results from the women's health initiative. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 34, 360.
- Quadrelli, S., Solís, M., 2018. *Personalized Lung Cancer Treatment: A Teamwork*.
- Relin Yang, M.D., Cheung, M.C., Byrne, M.M., Youjie Huang, M.D., Dao Nguyen, M.D., Lally, B.E., Koniaris, L.G., 2010. Do racial or socioeconomic disparities exist in lung cancer treatment? *Cancer* 116, 2437–2447.
- Román, M., Baraibar, I., López, I., Nadal, E., Rolfo, C., Vicent, S., Gilbazo, I., 2018. KRAS oncogene in non-small cell lung cancer: clinical perspectives on the treatment of an old target. *Mol. Cancer* 17, 33.
- Ryn, M.V., 2002. Research on the provider contribution to race/ethnicity disparities in medical care. *Medical Care* 40, 1140.
- Schabath, M.B., Cress, D., Munoz-Antonia, T., 2016a. Racial and ethnic differences in the epidemiology and genomics of lung cancer. *Cancer Control* 23, 338–346.
- Schabath, M.B., Cress, W.D., Muñozantonia, T., 2016b. Racial and Ethnic Differences in the Epidemiology of Lung Cancer and the Lung Cancer Genome.
- Soo, R.A., Kawaguchi, T., Loh, M., Ou, S.H., Shieh, M.P., Cho, B.C., Mok, T.S., Soong, R., 2012. Differences in outcome and toxicity between Asian and caucasian patients with lung cancer treated with systemic therapy. *Future Oncol.* 8, 451.
- Yu, E., Davidson, S., 2018. Current overview of lung cancer from pathology, screening to all treatment modalities. *Curr. Cancer Therap. Rev.* 14.
- Zhang, W., Edwards, A., Flemington, E.K., Zhang, K., 2017. Racial disparities in patient survival and tumor mutation burden, and the association between tumor mutation burden and cancer incidence rate. *Sci. Rep.* 7.