EBioMedicine 42 (2019) 16–17

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

EBioMedicine
Published by THE LANCET

Commentary

Premalignant genomic data tracing the evolution of lung adenocarcinoma

Cheng Wang [a,b,c], Hongbing Shen [a,b,*]

[a] Department of Epidemiology and Biostatistics, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, China
[b] Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Jiangsu Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing, China
[c] Department of Bioinformatics, School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing, China

Keywords:
Lung cancer

By analyzing SNP array data on "trios" samples including premalignant atypical adenomatous hyperplasias (AAH) of the lung, matched normal and lung adenocarcinoma (LUAD) tissues from 16 patients, Sivakumar et al. [1] in this issue investigated the allelic imbalance (AI) events in the premalignant lesions of lung adenocarcinoma. With a modified statistic method to infer the AI events sensitively, they identified AI events in more than 50% of AAH tissues, indicating the importance of chromosomal aberrations and genome instability in the development of AAH. In addition, they identified a number of chromosomal aberrations harboring tumor suppressors and oncogenes, such as 17p loss affecting well-known *TP53*. It was worth to note that the same group previously reported the somatic mutations and RNA-Seq data from the same samples [2]. Thus, they presented a comprehensive molecular landscape of the evolution process from AAH to LUAD, which is important for understanding the etiology of LUAD.

First of all, the premalignant data provided evidence that chromosomal aberrations initiated at the early stage of LUAD and emerged even earlier than many driver mutations. Both somatic mutations and copy number aberrations were the hallmarks of cancer. The TCGA study revealed a heavy burden of somatic mutations and copy number alterations in LUAD and most of them are "passengers" that have no effect on the neoplastic process [3]. Thus, it is difficult to identify which somatic alterations are drivers that promote a selective growth advantage. Previous studies suggested that genes carried "driver" alterations had a higher alteration rate than other genes and proposed statistical methods to identify driver genes [4]. By using the methods, TCGA studies reported that *TP53* was a commonly mutated (46%) gene in LUAD [5]. In addition, somatic alterations usually affect oncogenes in the receptor tyrosine kinase (RTK)/RAS/RAF pathways in LUAD and tumor suppressor genes including *STK11*, *KEAP1*, *NF1*, *RB1*, and *CDKN2A* [5]. They also observed copy number amplified regions (3q26, 5p15, 7p11, 8q24, 11q13, 12q13, 12q15, 14q13) and deleted regions (17p13, 9p21, 9p23) in LUAD [5]. However, it is difficult to decipher the causal relationship between somatic mutations and copy number alterations in the cancer genomics data. With the evidence from AAH, we can differentiate the initial driver alterations of LUAD like 17p loss involving *TP53*.

Furthermore, in this issue Sivakumar et al. [1] reported an important association between AI burden and smoking behavior in AAH. Interestingly, a similar trend was observed in their matched LUAD samples, but the difference was not significant, and at the same time, a higher AI burden was seen in *EGFR* mutated non-smokers. The results suggested that the underlying mechanisms of AI were quite different in smokers and non-smokers: AI served as a founding event in smoking patients and initiated the development of LUAD, while it was only a bypass product of genome instability in non-smoker patients and therefore occurred only in LUAD tissues. In the phylogenetic tree construction, Sivakumar et al. [1] found that 6 of 7 patients with evidence of shared AI between matched AAH and LUAD were smokers, which further supported the conclusion above. Thus, the AI events in smokers and non-smokers should not be considered equally in the following studies. However, the sample size involved in this analysis is relatively small and more well-designed epidemiological studies with larger sample size are warranted to further illuminate the differences.

The molecular landscapes of premalignant tissues could also be used to identify potential biomarkers to predict the progress from the prelesions to cancers. Teiseira et al. recently profiled the genomic, transcriptomic, and epigenomic landscape of the prelesion of lung squamous cancer [6]. With longitudinally monitoring information, they successfully generated a predictive model to identify which lesions will progress with remarkable accuracy, which offered a new strategy for the early detection of lung cancer.

Lastly, the public available multi-omics data from "trios" LUAD patients are a valuable resource for the following studies of LUAD. For instance, we recently conducted whole-genome sequencing and

DOI of original article: https://doi.org/10.1016/j.ebiom.2019.03.020.
* Corresponding author at: Department of Epidemiology and Biostatistics, Center for Global Health, School of Public Health, Nanjing Medical University, 101 Longmian Avenue, Nanjing 211166, China.
E-mail address: hbshen@njmu.edu.cn (H. Shen).

https://doi.org/10.1016/j.ebiom.2019.03.065

RNA-Seq on 90 lung cancer and observed an association between infiltrating B cells and *EGFR* mutations in tumor tissues of LUAD [7]. We performed a re-analysis of the RNA-Seq data from the previous Sivakumar et al.'s paper [2] and found that the elevation of infiltrating B cells and inflammatory microenvironments preceded the emergence of *EGFR* mutations [7].

However, there are still great challenges to investigate the evolution process from AAH to LUAD. Intratumor heterogeneity of lung cancer was one of them. A recent study conducted whole-exome sequencing on the multi-regions of non-small-cell lung cancer (NSCLC) [8]. They observed widespread intratumor heterogeneity for somatic copy-number alterations as well as mutations. A number of known driver alterations occurred in the sub-clones of LUAD. To clearly infer the clonal evolution and trace the origin of these alterations, a multi-region sampling of AAH are necessary for future studies. In addition, the low cell fraction of somatic alterations in AAH also limited the genomic studies and the subsequent evolution analysis. The authors overcome the challenge by jointly modeling the allelic imbalance statistics, but the technology advance may provide new insights into the era. Single-cell RNA-Seq has recently used to decipher the microenvironments and subclones of lung cancers [9] and additional methods have been developed to estimate the copy number alterations and copy-neutral loss-of-heterozygosity events in single-cell RNA-Seq data [10], which might be used in the future studies of AAH to solve the problems of cell heterogeneity.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Sivakumar S, FAS Lucas, Jakubek YA, TL McDowell, Lang W, et al. Genomic landscape of allelic imbalance in premalignant atypical adenomatous hyperplasias of the lung. EBioMedicine 2019. https://doi.org/10.1016/j.ebiom.2019.03.020.
[2] Sivakumar S, Lucas FAS, McDowell TL, Lang W, Xu L, Fujimoto J, et al. Genomic landscape of atypical adenomatous hyperplasia reveals divergent modes to lung adenocarcinoma. Cancer Res 2017;77(22):6119–30.
[3] Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz Jr LA, Kinzler KW. Cancer genome landscapes. Science 2013;339(6127):1546–58.
[4] Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature 2013;499(7457):214–8.
[5] Cancer Genome Atlas Research N. Comprehensive molecular profiling of lung adenocarcinoma. Nature 2014;511(7511):543–50.
[6] Teixeira VH, Pipinikas CP, Pennycuick A, Lee-Six H, Chandrasekharan D, Beane J, et al. Deciphering the genomic, epigenomic, and transcriptomic landscapes of pre-invasive lung cancer lesions. Nat Med 2019;25(3):517–25.
[7] Wang C, Yin R, Dai J, Gu Y, Cui S, Ma H, et al. Whole-genome sequencing reveals genomic signatures associated with the inflammatory microenvironments in Chinese NSCLC patients. Nat Commun 2018;9(1):2054.
[8] Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TBK, Veeriah S, et al. Tracking the evolution of non-small-cell lung cancer. N Engl J Med 2017;376(22):2109–21.
[9] Lambrechts D, Wauters E, Boeckx B, Aibar S, Nittner D, Burton O, et al. Phenotype molding of stromal cells in the lung tumor microenvironment. Nat Med 2018;24(8):1277–89.
[10] Fan J, Lee HO, Lee S, Ryu DE, Lee S, Xue C, et al. Linking transcriptional and genetic tumor heterogeneity through allele analysis of single-cell RNA-seq data. Genome Res 2018;28(8):1217–27.