Online Supplementary Material

H. Bai, J. Duan, C. Li et al. EPHA mutation as potential predictor to immunotherapeutic efficacy in lung adenocarcinoma

Corresponding author:

Jie Wang, MD, PhD, State Key Laboratory of Molecular Oncology, Department of Medical Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, 100021, China. 17 Pan-jia-yuan South Lane, Chaoyang District, 100021, Beijing, China. zlhuxi@163.com; Tel: 86-13910704669; Fax: 86-10-87788027.

Zhijie Wang, MD, State Key Laboratory of Molecular Oncology, Department of Medical Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China. 17 Pan-jia-yuan South Lane, Chaoyang District, 100021, Beijing, China. Tel: 86-10-87788029; Email: Jie_969@163.com; Fax: 86-87788027.

Dr. Li Zhang, Sun Yat-sen University Cancer Center, 651# East Dong Feng Road, Guangzhou, Guangdong, 510060, China. Email: zhangli6@mail.sysu.edu.cn

Supplemental methods and reference

Whole exome sequencing (WES) and predicted neoantigens

WES were implemented in China cohort according to the method that had been previously published [1]. Briefly, genomic DNAs from FFPE sections or biopsy samples and the whole blood control samples were extracted with QIAamp DNA FFPE Tissue Kit and DNeasy Blood and tissue kit (Qiagen, USA), respectively, and quantified by Qubit 3.0 using the dsDNA HS Assay Kit (ThermoFisher Scientific, USA). Library preparations were performed with KAPA Hyper Prep Kit (KAPA Biosystems, USA). Target enrichment was performed using the xGen Exome Research Panel and Hybridization and Wash Reagents Kit (Integrated DNA Technology, USA) according to manufacturer's protocol. Sequencing was performed on Illumina HiSeq4000 platform using PE150 sequencing chemistry (Illumina, USA). The average coverage size of WES for TMB estimation was 32 Mb. TMB was defined as the total number of nonsynonymous mutations. Neoantigens were predicted by *in silico* and obtained from the previous published paper

PD-L1 expression evaluation

PD-L1 immunohistochemistry (IHC) staining by SP263 (Ventana Medical Systems, Tucson, AZ) in a CLIA-accredited/CAP-certified laboratory was analyzed. PL-L1 expression was determined by tumor proportion score (TPS), which was defined as the percentage of tumor cells with partial or complete cell membrane staining at any intensity. A minimum of 100 evaluable tumor cells were required for determination of PD-L1 expression.

mRNA expression profiling analysis

Associations between *EPHA* mutation and the expression of genes involved in *TGF-β* signaling and immune signatures were analyzed in LUAD, LUSC, bladder, esophageal adenocarcinoma, skin cutaneous melanoma, and head and neck carcinoma in TCGA respectively, for whom both RNAseq and WES data were available. The immune gene list was mainly based on the published article that summarized the genes related to activated T cells, immune cytolytic activity and IFN-γ release [2-4]. The 43 core genes that regulate or mediate *TGF-β* signaling have been identified in previous study [5,6] and are available at cBioPortal (http://www.cbioportal.org) as "General: *TGF-β* superfamily". A list of 47 immune-related genes and 43 *TGF-β* genes are provided in **eTable 2 and eTable 3**. The different tumor cohorts for analysis with available RNAseq and DNAseq are also listed in **eTable 1**. The mRNA expression from cBioPortal was transformed by Z-score.

Gene set enrichment analysis (GSEA)

GSEA was performed using the java GSEA 3.0 Desktop Application (http://software.broadinstitute.org/gsea/index.jsp) [7,8] to identify whether immune signaling and *TGF-β* signaling genes were associated with *EPHA* status. The genes identified to be on the leading edge of the enrichment profile were subjected to pathway analyses. Genes with unavailable expression in more than 80% of samples were excluded from the GSEA analysis. The normalized enrichment score (NES) is the primary statistic for examining gene set enrichment results. The P value adjusted by FDR estimates the statistical significance of the enrichment score. A gene set with FDR≤0.05 was considered to be significantly enriched in genes.

Reference

1. Fang W, Ma Y, Yin JC, Hong S, Zhou H. Comprehensive Genomic Profiling Identifies Novel Genetic Predictors of Response to Anti-PD-(L)1 Therapies in Non-Small Cell Lung Cancer. 2019;25(16):5015-5026.
2. Dong ZY, Zhong WZ, Zhang XC, et al. Potential Predictive Value of and Mutation Status for Response to PD-1 Blockade Immunotherapy in Lung Adenocarcinoma. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2017;23(12):3012-3024.
3. Ayers M, Lunceford J, Nebozhyn M, et al. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *The Journal of clinical investigation.* 2017;127(8):2930-2940.
4. Wang Z, Zhao J, Wang G, et al. Comutations in DNA Damage Response Pathways Serve as Potential Biomarkers for Immune Checkpoint Blockade. *Cancer research.* 2018;78(22):6486-6496.
5. Korkut A, Zaidi S, Kanchi RS, et al. A Pan-Cancer Analysis Reveals High-Frequency Genetic Alterations in Mediators of Signaling by the TGF-β Superfamily. *Cell systems.* 2018;7(4):422-437.e427.
6. Gu S, Zaidi S, Hassan MI, et al. Mutated CEACAMs Disrupt Transforming Growth Factor Beta Signaling and Alter the Intestinal Microbiome to Promote Colorectal Carcinogenesis. *Gastroenterology.* 2019.
7. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences.* 2005;102(43):15545-15550.

8. Mootha VK, Lindgren CM, Eriksson K-F, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature genetics.* 2003;34(3):267-273.