



# Robust Prediction of Immune Checkpoint Inhibition Therapy for Non-Small Cell Lung Cancer

Jiehan Jiang<sup>1†</sup>, Zheng Jin<sup>2†</sup>, Yiqun Zhang<sup>2†</sup>, Ling Peng<sup>3</sup>, Yue Zhang<sup>4</sup>, Zhiruo Zhu<sup>5</sup>, Yaohui Wang<sup>5</sup>, De Tong<sup>5</sup>, Yining Yang<sup>2</sup>, Jianfei Wang<sup>2</sup>, Yadong Yang<sup>2</sup> and Kui Xiao<sup>5\*</sup>

<sup>1</sup> Department of Pulmonary and Critical Care Medicine, University of South China Affiliated Changsha Central Hospital, Changsha, China, <sup>2</sup> Research Institute, GloriousMed Clinical Laboratory (Shanghai) Co., Ltd, Shanghai, China, <sup>3</sup> Department of Respiratory Disease, Zhejiang Provincial People's Hospital, Hangzhou, China, <sup>4</sup> Tongji Medical College, Huazhong University of Science & Technology, Wuhan, China, <sup>5</sup> Department of Pulmonary and Critical Care Medicine, The Second Xiangya Hospital, Central South University, Changsha, China

## OPEN ACCESS

### Edited by:

Tao Jiang,  
Shanghai Pulmonary Hospital, China

### Reviewed by:

Hui Zheng,  
University of Science and Technology  
of China, China  
Yang Xia,  
Zhejiang University, China

### \*Correspondence:

Kui Xiao  
dr.kuixiao@csu.edu.cn

<sup>†</sup>These authors have contributed  
equally to this work and share  
first authorship

### Specialty section:

This article was submitted to  
Cancer Immunity and Immunotherapy,  
a section of the journal  
Frontiers in Immunology

**Received:** 28 December 2020

**Accepted:** 09 March 2021

**Published:** 13 April 2021

### Citation:

Jiang J, Jin Z, Zhang Y, Peng L,  
Zhang Y, Zhu Z, Wang Y, Tong D,  
Yang Y, Wang J, Yang Y and Xiao K  
(2021) Robust Prediction of Immune  
Checkpoint Inhibition Therapy for  
Non-Small Cell Lung Cancer.  
Front. Immunol. 12:646874.  
doi: 10.3389/fimmu.2021.646874

**Background:** The development of immune checkpoint inhibitors (ICIs) is a revolutionary milestone in the field of immune-oncology. However, the low response rate is the major problem of ICI treatment. The recent studies showed that response rate to single-agent programmed cell death protein 1 (PD-1)/programmed cell death-ligand 1 (PD-L1) inhibition in unselected non-small cell lung cancer (NSCLC) patients is 25% so that researchers defined several biomarkers to predict the response of immunotherapy in ICIs treatment. Common biomarkers like tumor mutational burden (TMB) and PD-L1 expression have several limitations, such as low accuracy and inadequately validated cutoff value.

**Methods:** Two published and an unpublished ICIs treatment NSCLC cohorts with 129 patients were collected and divided into a training cohort (n = 53), a validation cohort (n = 22), and two independent test cohorts (n = 34 and n = 20). We identified six immune-related pathways whose mutational status was significantly associated with overall survival after ICIs treatment. Then these pathways mutational status combined with TMB, PD-L1 expression and intratumor heterogeneity were incorporated to build a Bayesian-regularization neural networks (BRNN) model to predict the ICIs treatment response.

**Results:** We firstly proved that TMB, PD-L1, and mutant-allele tumor heterogeneity (MATH) were independent biomarkers. The survival analysis of six immune-related pathways revealed the mutational status could distinguish overall survival after ICIs treatment. When predicting immunotherapy efficacy, the overall accuracy of area under curve (AUC) in validation cohort reaches 0.85, outperforming previous predictors in either sensitivity or specificity. And the AUC in two independent test cohorts reach 0.74 and 0.80.

**Conclusion:** We developed a pathway-model that could predict the efficacy of ICIs in NSCLC patients. Our study made a significant contribution to solving the low prediction accuracy of immunotherapy of single biomarker. With the accumulation of larger data sets, further studies are warranted to refine the predictive performance of the approach.

**Keywords:** immunotherapy, non-small cell lung cancer, biomarkers, immune pathway, neural network, prognosis

## INTRODUCTION

Immunotherapy is emerging as a beneficial tool for cancer treatment by activating the immune system to produce antitumor effects (1). Recently, the most advanced approach to therapeutically utilize the antitumor activity is via immune checkpoint inhibitors (ICIs) (2). Immune checkpoint inhibitors work by releasing a natural brake on patient's immune system so that immune cells called T cells to recognize and attack tumors (3). Among the ICIs, programmed cell death protein 1 (PD-1)/programmed cell death-ligand 1 (PD-L1) and cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) inhibitors showed promising therapeutic outcomes, and some have been approved for numerous cancer therapy, such as melanoma, renal cell carcinoma (RCC), and non-small cell lung cancer (NSCLC) (4, 5). However, ICIs are not universally effective for all patients, and many patients fail to respond to ICIs due to intrinsic resistance or have an initial response followed by disease progression due to acquired resistance (6). For example, response rates to single-agent PD-1/PD-L1 inhibition in unselected patients with melanoma, NSCLC, and RCC are 40% (7, 8), 25% (8, 9), and 19% (10), respectively (11). To identify patients who are more likely to respond to PD-1/PD-L1 blockade as well as other immunotherapeutics, researchers defined several biomarkers to predict the response of immunotherapy in cancer treatment. The commonly used biomarkers include tumor mutational burden (TMB) and PD-L1 expression (11, 12). Patients with a higher TMB or higher PD-L1 expression have a higher likelihood of immunotherapy response. Another novel statistical value, mutant-allele tumor heterogeneity (MATH), has been documented that is not only as a measure of intratumor genetic heterogeneity but also can be used as a biomarker to predict the response of treatment for patients (13–16). In addition, recent studies have shown that some pathways, such as IFN- $\gamma$ , NF- $\kappa$ B, and Wnt, are cancer-related immune-regulation pathway, which may be potential indicators to explore the effect of immunotherapy (17–20).

Nevertheless, it has been documented that the available biomarkers have several limitations (21, 22), such as low accuracy, and inadequately validated cutoff value, and previous studies only use one or two of them independently in immunotherapy prediction (23). Therefore, we developed a pathway-model that included TMB, PD-L1, MATH, and immune-related pathway to predict the efficiency of ICIs, especially in NSCLC, which is the leading cause of cancer-related mortality worldwide (24). The pathway-model did not only have a high accuracy in published cohorts but also be proven to have an effective prediction ability in GloriousMed cohort with 20 NSCLC patients. This study made a significant contribution to solving the low prediction accuracy of immunotherapy of single biomarker.

## MATERIALS AND METHODS

### GloriousMed Cohort

Twenty patients with non-small cell lung cancer treated with PD-1/PD-L1 inhibitors in The Second Xiangya Hospital, Central

South University who had genomic profiling of whole exome sequencing (WES) before treatment were included in our GloriousMed cohort (**Supplementary Table S1**).

TMB was defined as the total number of somatic mutations per exome in megabases. PD-L1 staining was evaluated centrally by IHC using 22C3 antibody and an automated staining procedure developed by Dako. The percentage of PD-L1 expression was scored by a qualified pathologist in samples with a minimum of 100 viable tumor cells.

Objective response was assessed by investigator-assessed RECIST 1.1 criteria every 6 weeks (two cycles of ICB administration). The complete response (CR) or partial response (PR) was considered as responders, whereas patients with stable disease (SD) or progressive disease (PD) were considered as non-responders.

All patients collection and usage were in accordance with the principles of the Declaration of Helsinki and approved by the Institution Review Board of The Second Xiangya Hospital, Central South University. The written informed consent for sample acquisition was obtained from all patients. All data were deidentified.

### Public Cohorts

Three independent public cohorts including Hellmann cohort (25), Rizvi cohort (26), and Samstein cohort (27) were also used in this study. The data for the three independent cohorts were retrieved from published articles (**Supplementary Table S2**). Hellmann cohort included 75 NSCLC patients treated with combined PD-1 and CTLA-4 blockade. Rizvi cohort included 34 NSCLC patients that treated with pembrolizumab. The Samstein cohort contained 1,662 patients received immunotherapy from 11 different cancers.

### WES Sequencing

DNA was extracted from FFPE-fixed tumor tissue using QIAamp DNA FFPE Tissue Kit (Qiagen), and Genomic DNA (gDNA) was extracted from white blood cells using the Blood Genomic DNA Mini Kit (Cwbio). Integrated DNA Technologies's xGen Exome Research Panel v1.0 according to the standard procedures (IDT) were used to capture whole exome. For each sample, 200 to 500 ng FFPE DNA or 500 ng gDNA was then used for library preparation and quantification guided by KAPA Hyper Prep protocols (KAPA). Libraries were then purified by AMPure XP (Beckman) and quantified by Qubit<sup>TM</sup> dsDNA HS Assay Kit (Thermo Fisher). Final library was sequenced on the Illumina Novoseq6000 (PE150). Sequencing adapters were trimmed by Trimmomatic from the raw data (28). The reads after adapter trimming were then aligned with the human reference genome hg19 by BWA (29). Duplicated reads were removed by Picard. Mapped reads were also realigned to the genome by Genome Analysis Tool Kit. Somatic mutations were called by Mutect2 with a paired workflow. Variants were then annotated by ANNOVAR and self-development code (30). An in-house script was used to verify the human identity concordance of paired samples. Somatic mutations were filtered with the following rules: (1) base quality value  $\geq 20$ ; (2) mutation reads depth  $\geq 10$ ; (3) variant allele frequency  $\geq 5\%$ ; (4) reads supporting variation  $< 4$  and frequency  $< 2\%$  in normal, tumor abundance/normal abundance  $\geq 8$ ;

(5) no strand bias (GATK parameter FS > 60 for SNP and FS >200 for indel); (6) discard synonymous mutations.

## Quantitative and Statistical Analyses

TMB and PD-L1 expression of Hellmann cohort and Rizvi cohort were retrieved from published articles. MATH was calculated through R package maftools for GloriousMed, Hellmann and Rizvi cohorts (31). Correlation among TMB, MATH, and PD-L1 expression (%) were examined by the Pearson rank correlation method. Correlation between TMB or MATH and grouped PD-L1 expression were examined by the Wilcoxon signed-rank test.

The overall survival (OS) was defined from the start of ICIs treatment until death due to any cause. And the progression-free survival (PFS) was defined as the time from the start of ICIs treatment until disease progression. Of notes, the Samstein cohort merely published OS data and Rizvi cohort provided PFS data. The Kaplan-Meier method was used to estimate OS or PFS, and the log-rank test was used to compare the survival curves. All tests with a  $p$  value  $\leq 0.05$  were considered statistically significant.

## Immune-Related Pathway Selection

The detailed profiles of genes involved in HRR, MMR, BER, JAK, MAPK, PI3K, NF- $\kappa$ B, and Wnt pathways were listed in **Supplementary Table S3**. At first, mutational status of aforementioned six immune-related pathways in every sample was classified into two categories: the first one assigned with 0 (no non-synonymous mutation) and the second with 1 (at least one non-synonymous mutation). Then, DDR pathway mutation status of each sample was classified into three groups based on the mutational status of HRR, MMR, and BER. “N” represented no mutation in HRR, MMR, or BER, “C” was stood for co-mutation between HRR and MMR or BER, and “S” was other cases. In addition, the mutational status of PI3K, JAK, and NF- $\kappa$ B were integrated as one variable by summing the mutational status.

## Model Construction

Three models were constructed, one model with TMB, PD-L1 expression, MATH, and immune-related pathways, called “pathway-model”; a second with TMB, PD-L1 expression, and MATH, called “tri-model”; the last one, called “bivariate-model”, with TMB and PD-L1 expression (**Table 1**). Both TMB and MATH were z-score normalized. PD-L1 expression was stratified as 0% (Z), 1%-49% (L),  $\geq 50\%$  (H), or unknown (N). And immune-related pathways were processed according to Immune-Related Pathway Selection. All of the models were trained via Bayesian Regularized Neural Networks (BRNN) algorithm using corresponding variables with 2 layers and default hyperparameters from R package caret (32), and the resampling method “boot” was used to choose the optimal model. The cutoff value of single-factor variable, TMB, PD-L1 expression and MATH was estimated by BRNN algorithm as well. Fifty-three patients of the Hellmann cohort were used as the training set, and remaining 22 patients were validation set. Rizvi cohort and GloriousMed cohort were processed as above description and were used as testing cohort.

## Model Performance Evaluation

Receiver operating characteristic (ROC) curves were constructed with the predictor estimated from each of the previous models and single-factor variables with roc function of R package pROC (33). Benefit probability of each patient was extracted from prediction results, and DCB/NDB information was provided by the cohorts. Differences between DCB and NDB with benefit probability were examined by the Wilcoxon signed-rank test.

## Comprehensive Analysis of TCGA LUAD and LUSC Cohorts

The clinical information, RNA expression, mutational status and protein array of The Cancer Genome Atlas Lung Adenocarcinoma (TCGA LUAD) and Lung Squamous Cell Carcinoma (LUSC) patients were retrieved from TCGA database. The patients with EGFR exon 18–21 mutations and ALK gene fusions were filtered to avoid make a disturbance for the analysis. In the signature score analysis, the expression of genes in a signature was normalized in the form of fragments per kilobase of exon model per million mapped fragments (FPKM). Then, a principal component analysis (PCA) was performed, and PC1 was extracted to serve as gene signature score (34). The 18 signatures and their gene sets were summarized from published papers (34–38). The significantly differential expression analysis was based on DESeq2 (39). The row counts of LUAD and LUSC patients were used as input for DESeq2. The differential expression genes were defined as the genes with absolutely  $\log_2\text{Foldchange} \geq 1$  and  $p\text{-value} \leq 0.05$ . The oncoplot of top 30 mutated genes were drawn by using R package maftools (31).

## RESULTS

### TMB, PD-L1 Expression, and MATH Are Independent Variables

The previous studies documented that higher TMB or PD-L1 expression correlated with better outcomes as compared with lower TMB or PD-L1 expression (11, 12, 25, 40). However, in 70 of 75 patients from Hellmann cohort who had all three biomarkers data, correlation between TMB and PD-L1 expression was not significant ( $R = -0.14$ ,  $p\text{-value} = 0.24$ ). TMB of some patients was more than 10 but PD-L1 expression was less than 25% (**Figure 1A**). The results might reveal the biomarkers were not consistent in response prediction of ICIs treatment. In the meantime, the novel biomarker MATH was not significantly correlated with PD-L1 expression ( $R = -0.2$ ,  $p\text{-value} = 0.099$ ) or TMB ( $R = 0.14$ ,  $p\text{-value} = 0.24$ ) as well (**Figures 1B, C**). We further explored the correlation between stratified PD-L1 expression and TMB or MATH by stratifying PD-L1 expression as 0% (Z), 1% to 49% (L),  $\geq 50\%$  (H), and unknown (N). Neither MATH nor TMB showed a significant difference with any PD-L1 expression groups (**Figures 1D, E**). The Rizvi and GloriousMed cohort showed the consistent correlation results as well (**Supplementary Figure 1**). This lack of correlation suggested that TMB, PD-L1 expression,

**TABLE 1 |** Models and variables.

Model	Variable
Bivariate-model	TMB and PD-L1 expression
Tri-model	TMB, PD-L1 expression and MATH
Pathway-model	TMB, PD-L1 expression, MATH and immune-related pathways

TMB, tumor mutational burden; PD-L1, programmed cell death-ligand 1; MATH, mutant-allele tumor heterogeneity.

and MATH are independent predictive measures of response to ICIs treatment, and a robust model should be constructed to unify these variables.

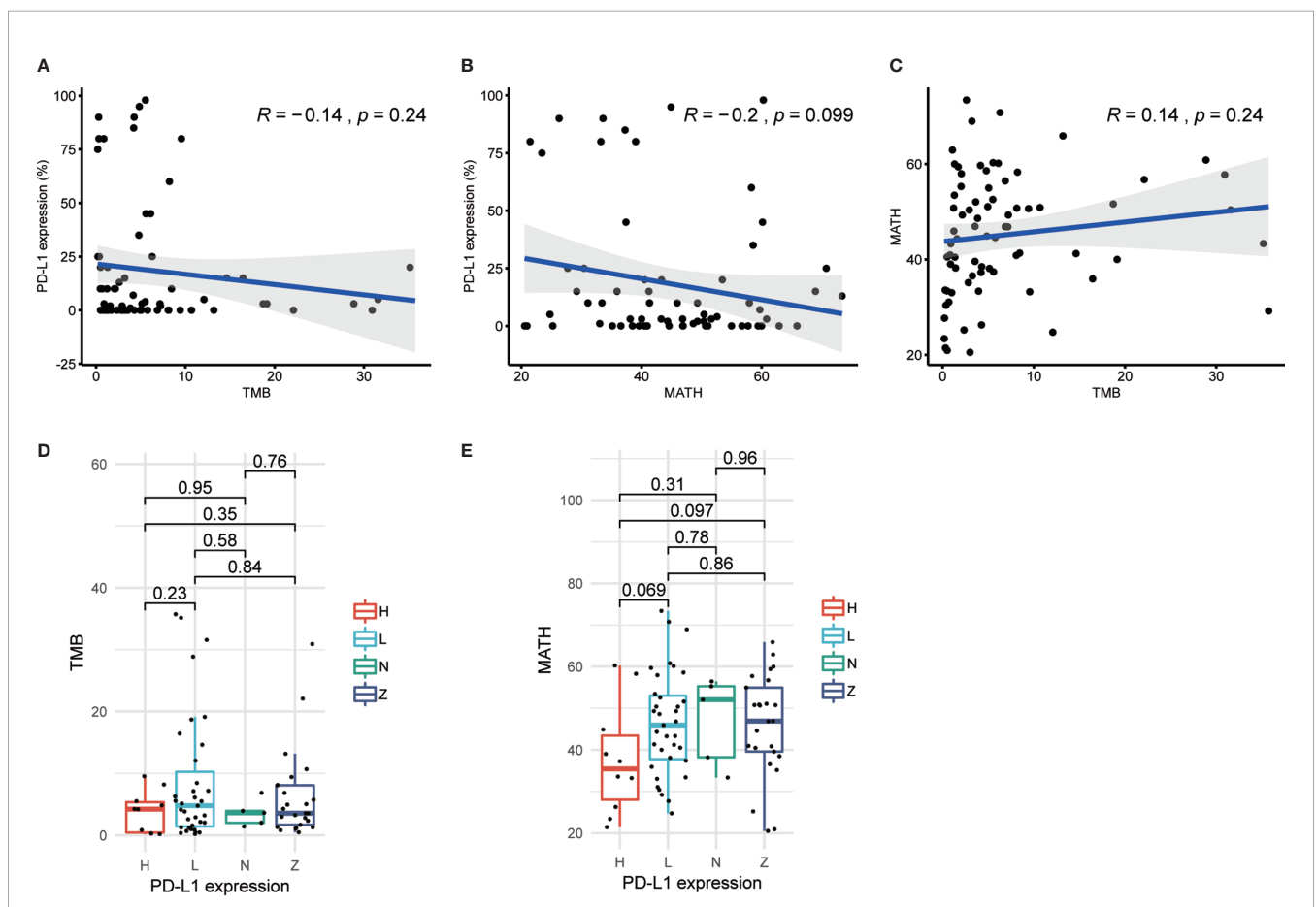
### Mutational Status of Immune-Related Pathway Can Act as Candidate Biomarkers

A prior study has shown that co-mutation information of DNA damage response (DDR) pathway can be used as a predictor of response to immune checkpoint blockade, and the mutation of the DDR solved the problem of difficulty in determining an

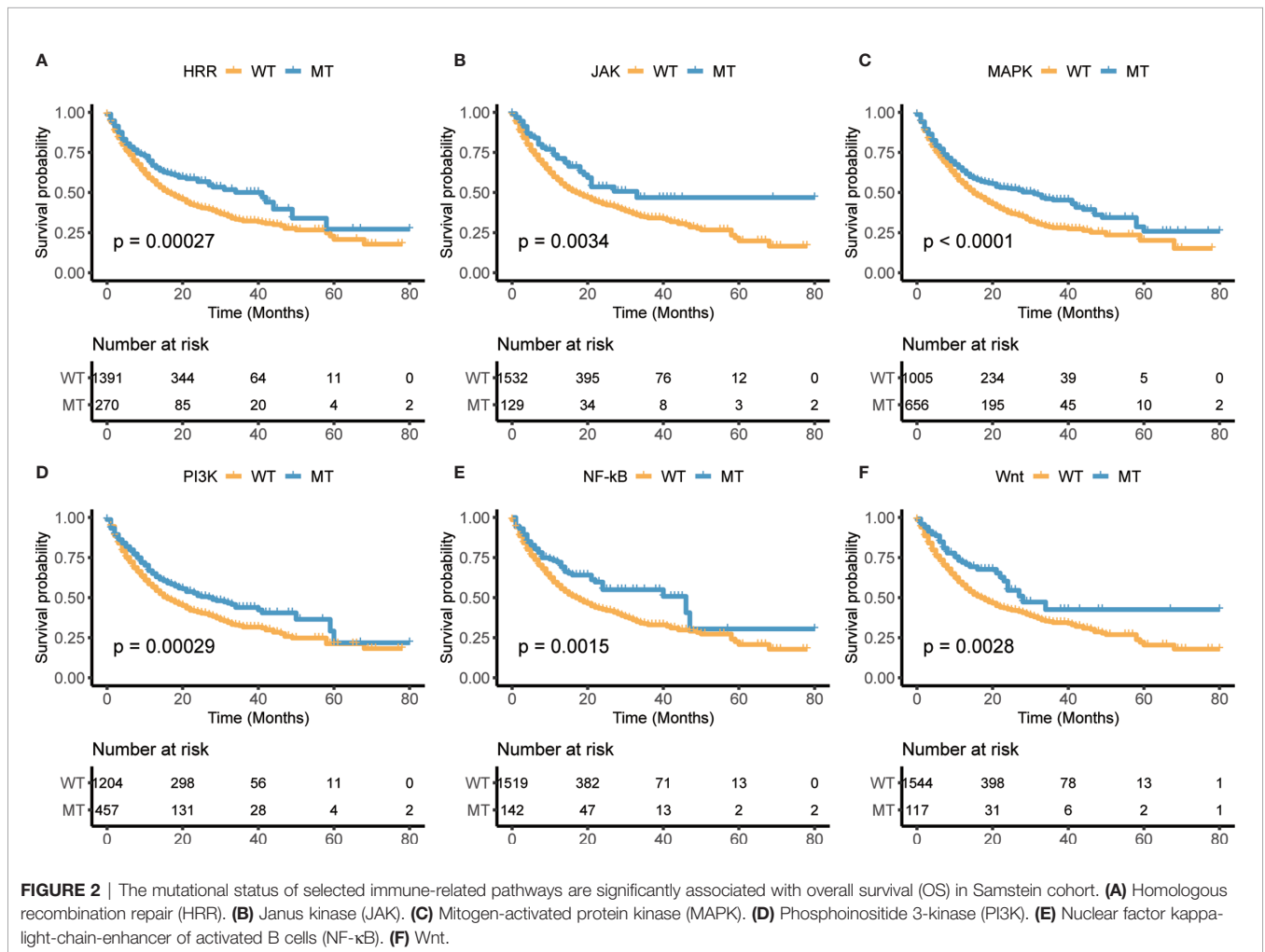
optimal TMB threshold (22). This finding provided a new way to predict the response of immunotherapy. Besides DDR pathway, we selected six pathways, homologous recombination repair (HRR), Janus kinase (JAK), mitogen-activated protein kinase (MAPK), phosphoinositide 3-kinase (PI3K), and nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB), Wnt, through literature survey, which are associated with tumor immunity or immunotherapy escape (41, 42). We also collected the mutational status of these pathways from Samstein cohort treated with ICIs (27) and explored its correlation with the overall survival (OS). The results showed that patients with mutations in any of six pathways had better survival than those without mutation (Figure 2). Furthermore, the results also revealed the selected pathways could be used as biomarkers to distinguish the prognosis for ICIs treatment.

### Pathway Model Is the Best Model to Predict the Efficiency of ICIs Treatment

We extracted 70% patients from Hellmann cohort, which totally included 75 NSCLC patients, as training data set (25) and the rest 30% patients were used to validate the models. Three different



**FIGURE 1 |** Tumor mutational burden (TMB), programmed cell death-ligand 1 (PD-L1) expression and mutant-allele tumor heterogeneity (MATH) are independent from each other in Hellmann cohort. (A) Scatterplot between TMB and PD-L1 expression (%). (B) Scatterplot between MATH and PD-L1 expression (%). (C) Scatterplot between TMB and MATH. (D) Boxplot of TMB and PD-L1 expression. (E) Boxplot of MATH and PD-L1 expression. The R value of (A–C) represents Pearson correlation coefficient.



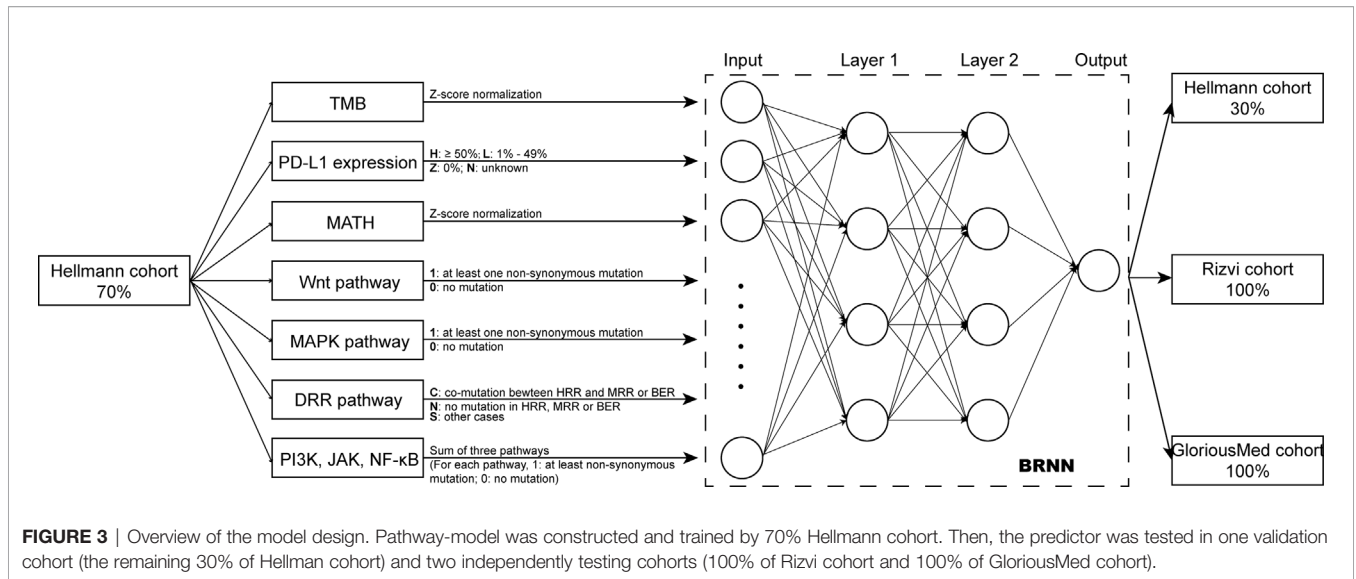
models were trained by using the training data set with different variables and were adjusted with clinical benefit as outcomes (Table 1). The pathway-model contains seven variables, including TMB, PD-L1 expression, MATH and the mutational status of six immune-related pathways (Figure 3). The mutational status of JAK, MAPK, and PI3K was integrated into one variable to improve the prediction accuracy. ROC curves based on the predictor for each of the three models estimated on Hellmann cohort (22 patients) were available and the results showed that the pathway-model was more predictive than other two models (AUC is 0.87, 0.83, and 0.59 for pathway-, tri-, and bivariate-model). The AUC of pathway-model was higher than single-factor variables containing TMB, PD-L1 expression, and MATH as well (AUC is 0.56, 0.49, and 0.69 for TMB, PD-L1 expression, and MATH) (Figure 4A and Table 2). We also checked the prediction benefit probability, a quantitative output generated from the model which represents the likelihood of immunotherapy response, of each patient compared with real clinical benefit information among three models. The benefit probability generated from pathway-model and tri-model are significantly higher in DCB group than in NDB group (p-value is 0.0024 for pathway-model and 0.0066 for tri-model), however, the median benefit probability of pathway-model (0.70) was higher

than tri-model (0.46). The difference of benefit probability was not significant in other models and single factors (Figure 4B).

We further tested the predictive ability of pathway-model in Rizvi cohort (26), consisting of 34 NSCLC patients treated with pembrolizumab, with all predictive variables and clinical benefit information available. The results showed that pathway-model could more accurately predict the clinical benefit of ICIs than other two models and single-factor variables (AUC is 0.74 for pathway-model, 0.67 for tri-model, 0.68 for bivariate-model, 0.63 for TMB, 0.72 for PD-L1 expression, and 0.55 for MATH) (Figure 4C and Table 2). The benefit probability of patients in DCB and NDB groups was significantly different as well (p-value is 0.0017, Figure 4D). The survival analysis indicated that the high benefit probability group also showed a better PFS (Figure 4E).

### Pathway Model Can Precisely Predict the Response of ICIs Treatment in GloriousMed Cohort

Finally, we tested pathway-model in GloriousMed cohort with 20 NSCLC patients, who were treated by PD-1/PD-L1 inhibitors (Supplementary Table S1). The accuracy of pathway-model was



much higher than tri-model and bivariate-model (AUC is 0.80 for pathway-model, 0.47 for tri-model and 0.64 for bivariate-model) (Figure 5A and Table 2). Even though, the benefit probability was not significantly different between DCB and NDB group (p-value is 0.08 for pathway-model), all DCB patients have a predictive benefit probability higher than 0.5 (Figure 5B). Thus, pathway-model can be generalized in clinical to improve the prediction accuracy of the response to immunotherapy.

### Comprehensive Analysis With TCGA NSCLC Cohort Imply that High Benefit Probability Patients Is Associated With Immune Response

We predicted the benefit probability of TCGA LUAD and TCGA LUSC cohorts without EGFR exon 18-21 mutations and ALK gene fusions patients in immunotherapy with pathway-model and classified patients to two groups at the median cut-point. Then, we calculated signature scores of 18 gene sets with principle component analysis (PCA) method. In TCGA LUAD cohort, thirteen signatures are significantly different between high benefit probability group and low probability group (Figure 6A). In consideration of TMB, and mutational status of DDR and Wnt pathways are included in prediction model, the benefit probability difference in DDR, WNT target, DNA repair-related signatures and cell cycle were expected. The signature score of CD 8 T effector and Immune Checkpoint were higher in high probability group than in that of low group, while the signature score of EMT3 and FGFR3 related was lower in high probability group (Figure 6A). However, in LUSC cohort, we did not find significant difference between high and low benefit probability groups as LUAD cohort (Figure 6D).

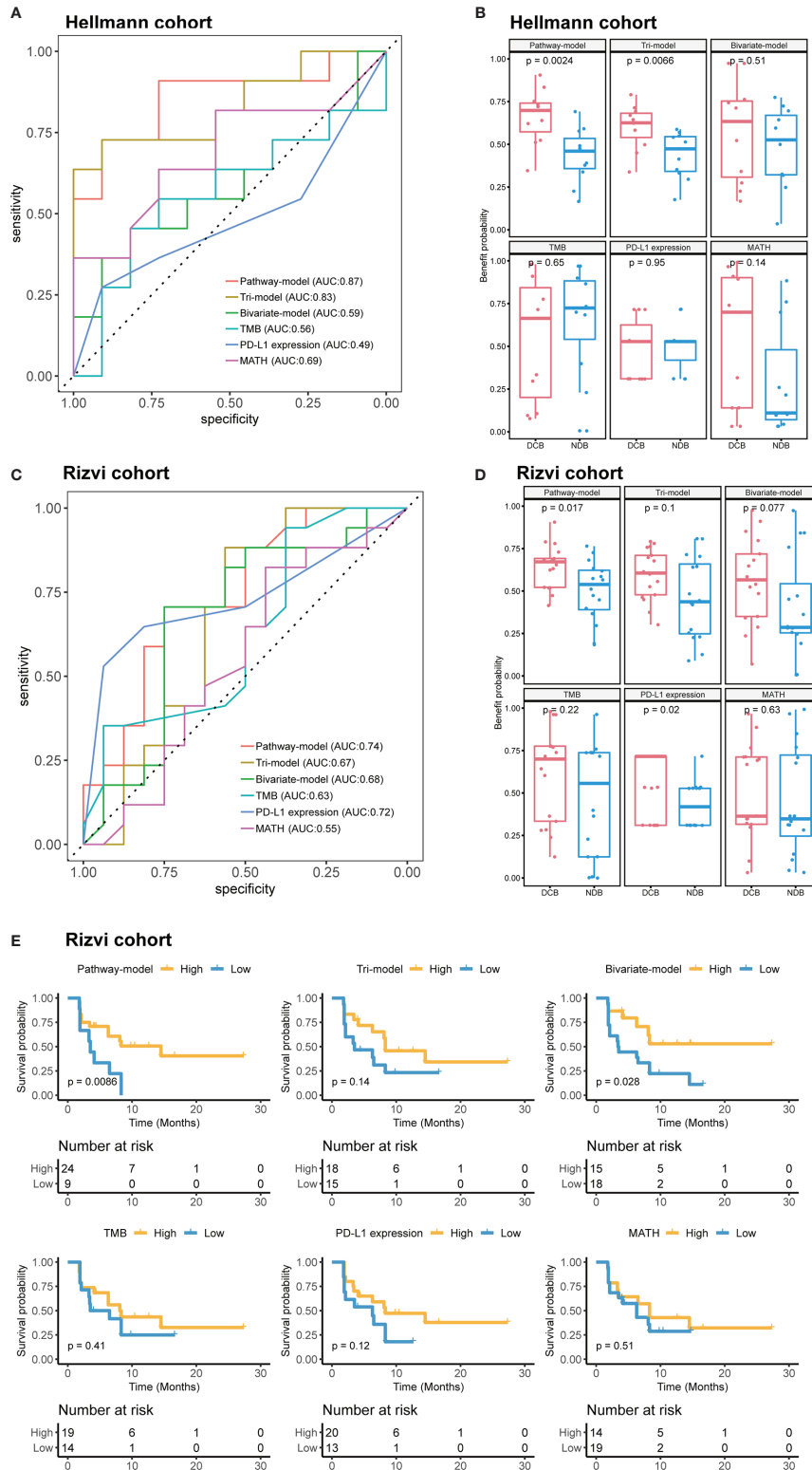
Furthermore, we analyzed the differential expression genes between high benefit probability groups and low group in LUAD and LUSC respectively (Figures 6B, E, Supplementary Table S4). There are 153 differential expression genes (106 up-regulated) in LUAD, including *AFP* and *G6PC*, which related

to P53 downstream pathway and FOXO pathway. In LUSC, there are 120 differential expression genes (50 up-regulated) including *FGF3* and *DLK1*, which related to FGFR pathway and NOTCH pathway. Apart from that, part of the top 30 mutated genes, such as *KRAS* and *PTPRD*, have different mutation pattern between high benefit probability group and low group, as well as between LUAD and LUSC (Figures 6C, F).

Above all, the comprehensive analysis of TCGA LUAD and LUSC cohorts imply that high benefit probability patients from pathway-model is associated with immune response.

### DISCUSSION

Immune checkpoints inhibitors (ICIs), such as PD-1 and PD-L1, have revolutionized the treatment of many cancers, including NSCLC. However, how to select patients most likely to benefit from immunotherapy is the current leading challenge in the field. Previous ICIs-related studies preferred to use several single biomarkers, respectively, to predict the prognosis of immunotherapy (25, 26). Our study constructed a robust pathway-model based on deep learning approach, which included two common biomarkers, TMB, PD-L1 expression, a recent developed intratumor heterogeneity evaluation value MATH and potential marker-immune-related pathways. To the best of our knowledge, this is the first study to combine mutational status of pathways and common biomarkers for efficacy of prediction in NSCLC. Not only the ROC curves but also the significant difference of benefit probability from our predictor between DCB and NDB showed that our model had high accuracy in both training and test NSCLC data sets. The comparison among our pathway-model, tri-model, bivariate-model, and single-factor variables showed that our pathway-model had the highest accuracy in predicting the response to ICIs treatment. We found that tri-model with MATH had a lower AUC than bivariate-model without MATH in Rizvi and



**FIGURE 4** | The performance comparison different models and single-factor variables of in validation cohort (Hellmann cohort) and independent test cohort (Rizvi cohort). **(A)** Receiver operating characteristic (ROC) curves of different models. **(B)** Benefit probability and risk of patients in different response groups. **(C)** ROC curves of different models. **(D)** Benefit probability and risk of patients in different response groups. **(E)** Survival analysis based on different models and single-factor variables, time was progression-free survival (PFS). Patients of **(A, B)** were from Hellmann cohort, and patients of **(C–E)** were from Rizvi cohort.

**TABLE 2** | Performance of models in three cohorts.

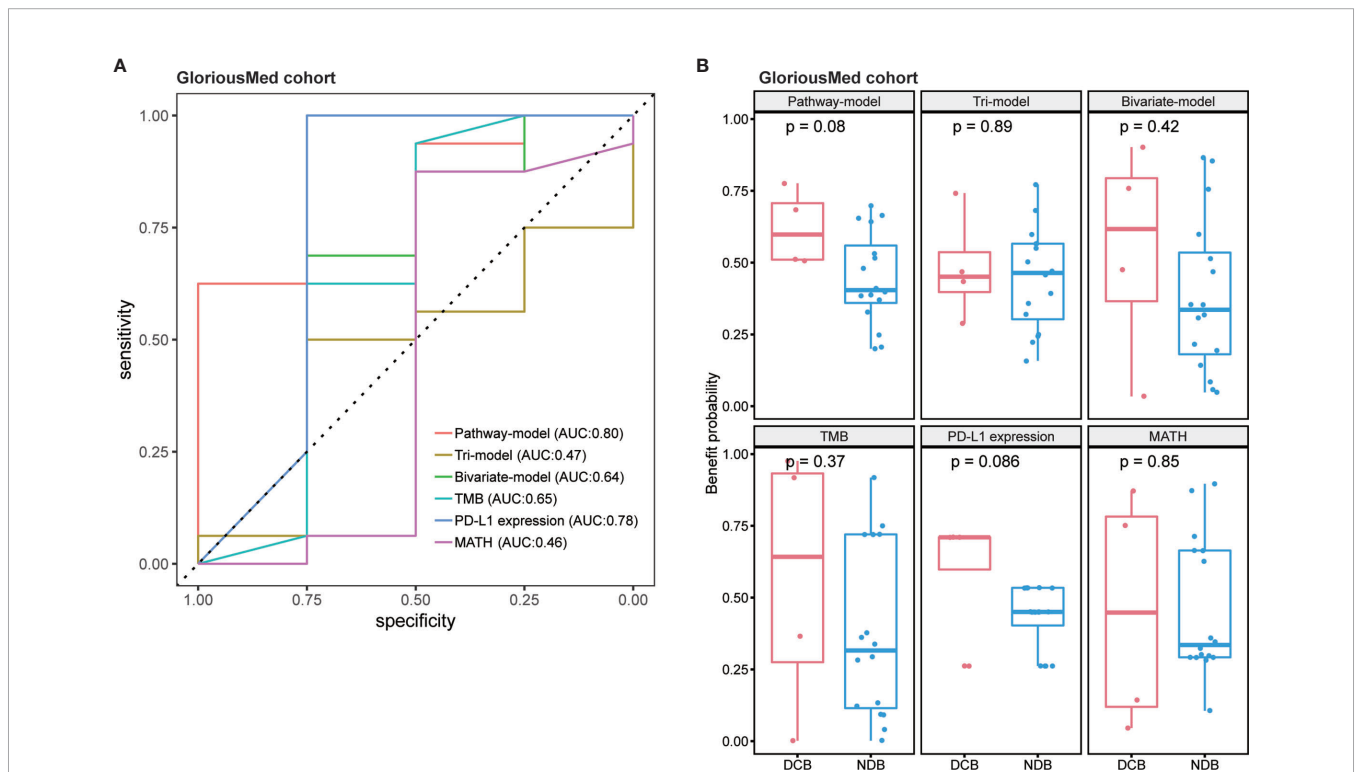
	Pathway-Model	Tri-model	Bivariate-model	TMB	PD-L1	MATH
Hellmann cohort	0.87	0.83	0.59	0.56	0.49	0.69
Rizvi cohort	0.74	0.67	0.68	0.63	0.72	0.55
GloriousMed cohort	0.80	0.47	0.64	0.65	0.78	0.46

TMB, tumor mutational burden; PD-L1, programmed cell death-ligand 1; MATH, mutant-allele tumor heterogeneity.

GloriousMed cohort. However, there is no denying that MATH did not improve the efficacy in distinguishing DCB and NDB patients in Rizvi and GloriousMed cohort in tri-model compared with bivariate-model. And pathway-model with MATH is the most stable model compared to other models and single factor variables. A recent study has shown that the integration of TMB and MATH forms a predictive marker for the response of ICIs treatment in melanoma (16), and another study has also revealed that intratumoral heterogeneity (MATH is an indicator of intratumoral heterogeneity) can be used as a biomarker to predict the response of ICIs treatment in NSCLC (15). Moreover, we found that the common biomarkers were not significant correlation according to the Pearson correlation coefficient, and the accuracy of each single-factor variable was lower than the pathway-model or tri-model. It might indicate

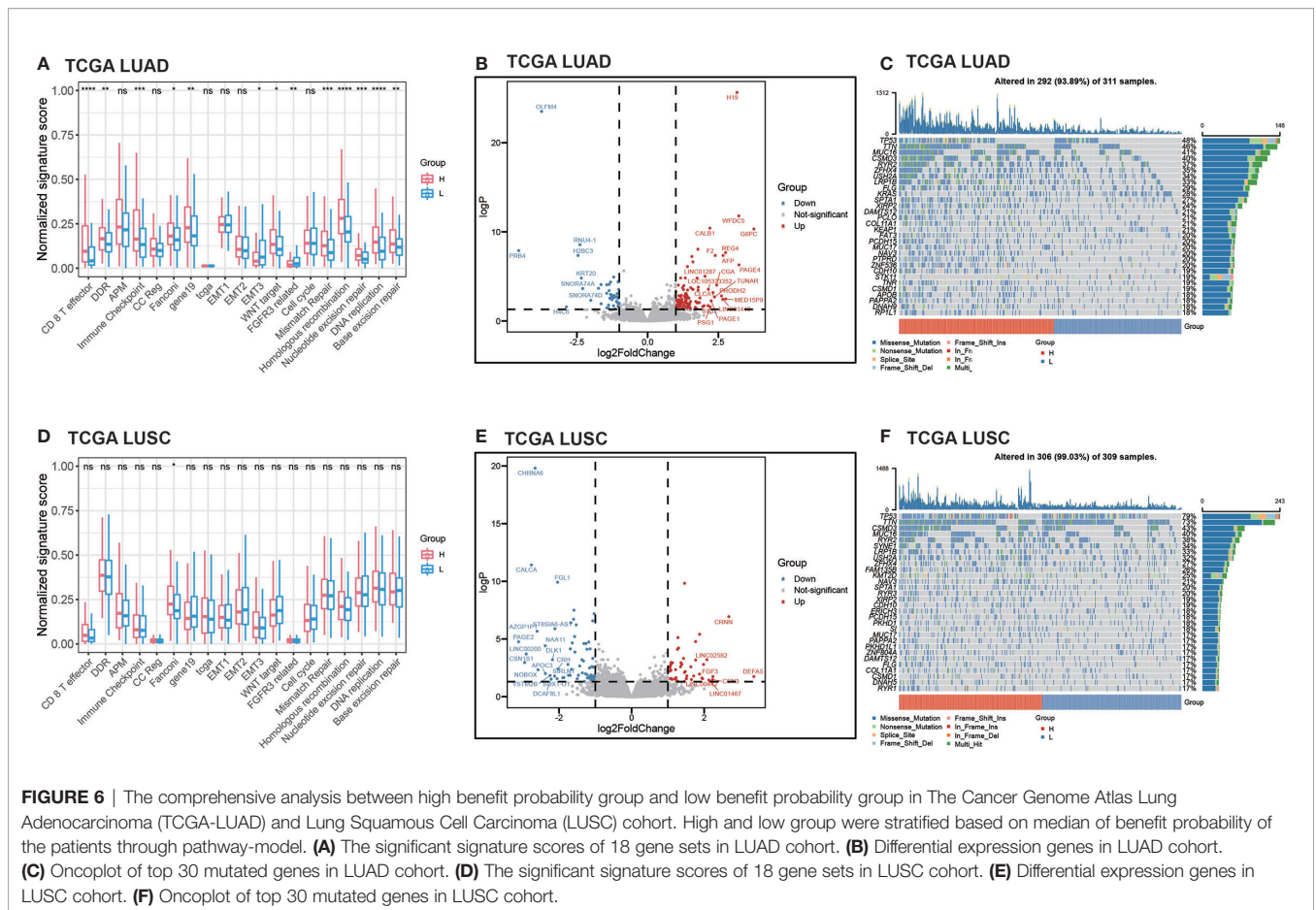
there was a great synergy among these biomarkers. When we grouped the patients at the median of benefit probability generated from pathway-model, the PFS time was significantly different between high and low group, specifically patients with high benefit probability were more likely to have longer PFS time. These results suggested that besides the ability of response prediction of ICIs treatment, benefit probability is also associated with the prognosis of NSCLC patients. In addition to, the prediction results of GloriousMed cohort prove that our pathway-model can effectively predict the benefit probability of ICIs treatment and can be generalized in clinical to provide some reference during the treatment.

Furthermore, the enrichment analysis of 18 immune-related gene sets in TCGA LUAD and LUSC cohort suggested that our model might reveal the possible mechanism of the immune phenotype of tumors. Previous studies have proven that CD8 cell play a central role in immunity to cancer through their capacity to kill malignant cells, EMT-related genes may contribute to tumor immune escape, and FGFR mutated cases have a more deserted immune phenotype than the wild type (43–46). Our immune infiltration analysis also showed that the high benefit probability group of LUAD cohort had higher CD8 T effector scores. However, the significant difference of signature scores between high benefit probability group and low group were only found in TCGA LUAD cohort, but not in TCGA LUSC cohort. It is implied that the underlying immune response



**FIGURE 5** | The performance comparison of different models in GloriousMed cohort. **(A)** Receiver operating characteristic (ROC) curves of different models. **(B)** Benefit probability and risk of patients in different response groups.





mechanism may be different between LUAD and LUSC. The differential expression genes in LUAD and LUSC are not complete same. P53 downstream pathway and FOXO pathway may be enriched in LUAD due to the up-regulation genes *AFP* and *G6PC*. P53 signaling pathway has been known as an important pathway in immune response, for example, it can function in immune cells including myeloid and T cells (47). Previous study has shown that FOXO pathway can be a target in tumor drug development (48). In LUSC, two differential expression genes, *FGF3* and *DLK1* are related two different pathways, FGFR pathway and Notch pathway. The enrichment of FGFR pathway implies a desert-immune subtype and high tumor purity of LUSC (45). Notch pathway can control the fate of various T cell type and myeloid cells that down-regulated *DLK1* might influence the immune cells (49). The different regulated pathways between LUAD and LUSC may be one of the reasons of different immune response mechanism. In LUAD cohort, the mutation ratio of *KRAS*, an oncogene which leads to immune escape in the tumor microenvironment (50), and *PTPRD*, which affects the tumor proliferation (51), were higher than LUSC also suggests the difference immune response mechanisms. All above inference is based on naïve treatment public cohort, the exact mechanism would still to be explored

with treatment samples. Except that, the probability of some differential expression genes, such as *MUC2*, *CLCA1*, *REG4*, and *FGF3* can be used as prognostic biomarkers in NSCLC is worth exploring because they have been reported as a biomarkers in other cancers as well (52–55).

There were limitations in our study that should be acknowledged. First, patients in the training cohort were treated with Nivolumab Plus Ipilimumab, and the model generated from which may be distracted in predicting patient in test cohort treated with Pembrolizumab or Tislelizumab due to pharmaceutical and medication differences. Second, the PD-L1 expression was quantified with different antibodies in training and validation cohort. Also, in the exploring cohort in TCGA data set, the PD-L1 expression was quantified using reverse phase protein array. The platform discordant of PD-L1 quantification may impair the power of our prediction model. Besides, due to the limitation of the training data sets, it is difficult to get a satisfactory model. Also, there are other features that are not incorporated into our model due to unavailability in either training or validation cohort, such as immune phenotype, which is known to affect the immunotherapy efficacy. In future studies, we will include more patients and features to guarantee the training process and the clinical practice of the predicting ICIs treatment efficacy in NSCLC patients.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Second Xiangya Hospital, Central South University. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

KX, YiY, JJ, and JW conceived and designed the project. YiZ, DT, and LP prepared and collected the data. ZJ, YuZ, YiZ, and ZZ contributed to analysis and interpretation. KX, ZJ, JJ, and YaY drafted the manuscript. KX, YW, JJ, JW, and YaY performed the quality assessment and revised the manuscript. All authors contributed to the article and approved the submitted version.

## REFERENCES

- Whiteside TL, Demaria S, Rodriguez-Ruiz ME, Zarour HM, Melero I. Emerging Opportunities and Challenges in Cancer Immunotherapy. *Clin Cancer Res* (2016) 22(8):1845–55. doi: 10.1158/1078-0432.CCR-16-0049
- Esfahani K, Roudaia L, Buhlaiga N, Del Rincon SV, Papneja N, Miller WH, Jr. A review of cancer immunotherapy: from the past, to the present, to the future. *Curr Oncol* (2020) 27(Suppl 2):S87–97. doi: 10.3747/co.27.5223
- Dyck L, Mills KHG. Immune checkpoints and their inhibition in cancer and infectious diseases. *Eur J Immunol* (2017) 47(5):765–79. doi: 10.1002/eji.201646875
- Darvin P, Toor SM, Sasidharan Nair V, Elkord E. Immune checkpoint inhibitors: recent progress and potential biomarkers. *Exp Mol Med* (2018) 50(12):1–11. doi: 10.1038/s12276-018-0191-1
- Park JA, Cheung NV. Limitations and opportunities for immune checkpoint inhibitors in pediatric malignancies. *Cancer Treat Rev* (2017) 58:22–33. doi: 10.1016/j.ctrv.2017.05.006
- Ottaviano M, De Placido S, Ascierto PA. Recent success and limitations of immune checkpoint inhibitors for cancer: a lesson from melanoma. *Virchows Arch* (2019) 474(4):421–32. doi: 10.1007/s00428-019-02538-4
- Robert C, Long GV, Brady B, Dutriaux C, Maio M, Mortier L, et al. Nivolumab in previously untreated melanoma without BRAF mutation. *N Engl J Med* (2015) 372(4):320–30. doi: 10.1056/NEJMoa1412082
- Weber JS, D'Angelo SP, Minor D, Hodi FS, Gutzmer R, Neyns B, et al. Nivolumab versus chemotherapy in patients with advanced melanoma who progressed after anti-CTLA-4 treatment (CheckMate 037): a randomised, controlled, open-label, phase 3 trial. *Lancet Oncol* (2015) 16(4):375–84. doi: 10.1016/S1470-2045(15)70076-8
- Garon EB, Rizvi NA, Hui R, Leigh N, Balmanoukian AS, Eder JP, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J Med* (2015) 372(21):2018–28. doi: 10.1056/NEJMoa1501824
- Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, et al. Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. *N Engl J Med* (2015) 373(19):1803–13. doi: 10.1056/NEJMoa1510665
- Goodman AM, Kato S, Bazhenova L, Patel SP, Frampton GM, Miller V, et al. Tumor Mutational Burden as an Independent Predictor of Response to Immunotherapy in Diverse Cancers. *Mol Cancer Ther* (2017) 16(11):2598–608. doi: 10.1158/1535-7163.MCT-17-0386
- Zhu J, Zhang T, Li J, Lin J, Liang W, Huang W, et al. Association Between Tumor Mutation Burden (TMB) and Outcomes of Cancer Patients Treated

## FUNDING

This study was supported by Scientific Research Project of Human Provincial Health Commission (No. 202103020704).

## ACKNOWLEDGMENTS

The authors would like to thank cBioPortal database, TCGA consortium and other involved research for their publicly available data. The results published here are in part based on data generated by the TCGA Research Network: <https://www.cancer.gov/tcga>.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2021.646874/full#supplementary-material>

- With PD-1/PD-L1 Inhibitions: A Meta-Analysis. *Front Pharmacol* (2019) 10:673. doi: 10.3389/fphar.2019.00673
- Greenbaum A, Martin DR, Bocklage T, Lee JH, Ness SA, Rajput A. Tumor Heterogeneity as a Predictor of Response to Neoadjuvant Chemotherapy in Locally Advanced Rectal Cancer. *Clin Colorectal Cancer* (2019) 18(2):102–9. doi: 10.1016/j.clcc.2019.02.003
  - Mroz EA, Rocco JW. MATH, a novel measure of intratumor genetic heterogeneity, is high in poor-outcome classes of head and neck squamous cell carcinoma. *Oral Oncol* (2013) 49(3):211–5. doi: 10.1016/j.oraloncology.2012.09.007
  - Fang W, Jin H, Zhou H, Hong S, Ma Y, Zhang Y, et al. Intratumoral heterogeneity as a predictive biomarker in anti-PD-(L)1 therapies for non-small cell lung cancer. *Mol Cancer* (2021) 20(1):37. doi: 10.1186/s12943-021-01331-9
  - Gao Y, Yang C, He N, Zhao G, Wang J, Yang Y. Integration of the Tumor Mutational Burden and Tumor Heterogeneity Identify an Immunological Subtype of Melanoma With Favorable Survival. *Front Oncol* (2020) 10:571545. doi: 10.3389/fonc.2020.571545
  - Pai SG, Carneiro BA, Mota JM, Costa R, Leite CA, Barroso-Sousa R, et al. Wnt/beta-catenin pathway: modulating anticancer immune response. *J Hematol Oncol* (2017) 10(1):101. doi: 10.1186/s13045-017-0471-6
  - Gattinoni L, Ji Y, Restifo NP. Wnt/beta-catenin signaling in T-cell immunity and cancer immunotherapy. *Clin Cancer Res* (2010) 16(19):4695–701. doi: 10.1158/1078-0432.CCR-10-0356
  - Ni L, Lu J. Interferon gamma in cancer immunotherapy. *Cancer Med* (2018) 7(9):4509–16. doi: 10.1002/cam4.1700
  - Xia L, Tan S, Zhou Y, Lin J, Wang H, Oyang L, et al. Role of the NFkappaB-signaling pathway in cancer. *Oncotargets Ther* (2018) 11:2063–73. doi: 10.2147/OTT.S161109
  - Berland L, Heeke S, Humbert O, Macocco A, Long-Mira E, Lassalle S, et al. Current views on tumor mutational burden in patients with non-small cell lung cancer treated by immune checkpoint inhibitors. *J Thorac Dis* (2019) 11(Suppl 1):S71–80. doi: 10.21037/jtd.2018.11.102
  - Wang Z, Zhao J, Wang G, Zhang F, Zhang Z, Zhang F, et al. Computations in DNA Damage Response Pathways Serve as Potential Biomarkers for Immune Checkpoint Blockade. *Cancer Res* (2018) 78(22):6486–96. doi: 10.1158/0008-5472.CAN-18-1814
  - Richard C, Fumet JD, Chevrier S, Derangere V, Ledys F, Lagrange A, et al. Exome Analysis Reveals Genomic Markers Associated with Better Efficacy of Nivolumab in Lung Cancer Patients. *Clin Cancer Res* (2019) 25(3):957–66. doi: 10.1158/1078-0432.CCR-18-1940

24. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca-Cancer J Clin* (2018) 68(6):394–424. doi: 10.3322/caac.21492
25. Hellmann MD, Nathanson T, Rizvi H, Creelan BC, Sanchez-Vega F, Ahuja A, et al. Genomic Features of Response to Combination Immunotherapy in Patients with Advanced Non-Small-Cell Lung Cancer. *Cancer Cell* (2018) 33(5):843–52 e4. doi: 10.1016/j.ccell.2018.03.018
26. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* (2015) 348(6230):124–8. doi: 10.1126/science.aaa1348
27. Samstein RM, Lee CH, Shoushtari AN, Hellmann MD, Shen R, Janjigian YY, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet* (2019) 51(2):202–6. doi: 10.1038/s41588-018-0312-8
28. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* (2014) 30(15):2114–20. doi: 10.1093/bioinformatics/btu170
29. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* (2009) 25(14):1754–60. doi: 10.1093/bioinformatics/btp324
30. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* (2010) 38(16):e164. doi: 10.1093/nar/gkq603
31. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* (2018) 28(11):1747–56. doi: 10.1101/gr.239244.118
32. Kuhn M. Building Predictive Models in R Using the caret Package. *J Stat Softw* (2008) 28(5):1–26. doi: 10.18637/jss.v028.i05
33. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S plus to analyze and compare ROC curves. *BMC Bioinf* (2011) 12(1):1–18. doi: 10.1186/1471-2105-12-77
34. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* (2018) 554(7693):544–8. doi: 10.1038/nature25501
35. Balar AV, Galsky MD, Rosenberg JE, Powles T, Petrylak DP, Bellmunt J, et al. Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial. *Lancet* (2017) 389(10064):67–76. doi: 10.1016/S0140-6736(16)32455-2
36. Cancer Genome Atlas Research N. Comprehensive molecular characterization of urothelial bladder carcinoma. *Nature* (2014) 507(7492):315–22. doi: 10.1038/nature12965
37. Lange SS, Takata K, Wood RD. DNA polymerases and cancer. *Nat Rev Cancer* (2011) 11(2):96–110. doi: 10.1038/nrc2998
38. Rosenberg JE, Hoffman-Censits J, Powles T, van der Heijden MS, Balar AV, Necchi A, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. *Lancet* (2016) 387(10031):1909–20. doi: 10.1016/S0140-6736(16)00561-4
39. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* (2014) 15(12):550. doi: 10.1186/s13059-014-0550-8
40. Zhao D, Mambetsariev I, Li H, Chen C, Fricke J, Fann P, et al. Association of molecular characteristics with survival in advanced non-small cell lung cancer patients treated with checkpoint inhibitors. *Lung Cancer* (2020) 146:174–81. doi: 10.1016/j.lungcan.2020.05.025
41. Bai J, Gao Z, Li X, Dong L, Han W, Nie J. Regulation of PD-1/PD-L1 pathway and resistance to PD-1/PD-L1 blockade. *Oncotarget* (2017) 8(66):110693–707. doi: 10.18632/oncotarget.22690
42. Hoessel B, Schmid JA. The complexity of NF-kappaB signaling in inflammation and cancer. *Mol Cancer* (2013) 12:86. doi: 10.1186/1476-4598-12-86
43. Durgeau A, Virk Y, Corgnac S, Mami-Chouaib F. Recent Advances in Targeting CD8 T-Cell Immunity for More Effective Cancer Immunotherapy. *Front Immunol* (2018) 9:14. doi: 10.3389/fimmu.2018.00014
44. Jackson SR, Yuan J, Teague RM. Targeting CD8+ T-cell tolerance for cancer immunotherapy. *Immunotherapy* (2014) 6(7):833–52. doi: 10.2217/imt.14.51
45. Zeng D, Ye Z, Wu J, Zhou R, Fan X, Wang G, et al. Macrophage correlates with immunophenotype and predicts anti-PD-L1 response of urothelial cancer. *Theranostics* (2020) 10(15):7002–14. doi: 10.7150/thno.46176
46. Terry S, Savagner P, Ortiz-Cuaran S, Mahjoubi L, Saintigny P, Thiery JP, et al. New insights into the role of EMT in tumor immune escape. *Mol Oncol* (2017) 11(7):824–46. doi: 10.1002/1878-0261.12093
47. Blagih J, Buck MD, Vousden KH. p53, cancer and the immune response. *J Cell Sci* (2020) 133(5). doi: 10.1242/jcs.237453
48. Farhan M, Wang H, Gaur U, Little PJ, Xu J, Zheng W. FOXO Signaling Pathways as Therapeutic Targets in Cancer. *Int J Biol Sci* (2017) 13(7):815–27. doi: 10.7150/ijbs.20052
49. Janghorban M, Xin L, Rosen JM, Zhang XH. Notch Signaling as a Regulator of the Tumor Immune Response: To Target or Not To Target? *Front Immunol* (2018) 9:1649. doi: 10.3389/fimmu.2018.01649
50. Hamarsheh S, Gross O, Brummer T, Zeiser R. Immune modulatory effects of oncogenic KRAS in cancer. *Nat Commun* (2020) 11(1):5439. doi: 10.1038/s41467-020-19288-6
51. Veeriah S, Brennan C, Meng S, Singh B, Fagin JA, Solit DB, et al. The tyrosine phosphatase PTPRD is a tumor suppressor that is frequently inactivated and mutated in glioblastoma and other human cancers. *Proc Natl Acad Sci USA* (2009) 106(23):9435–40. doi: 10.1073/pnas.0900571106
52. Li C, Zuo D, Yin L, Lin Y, Li C, Liu T, et al. Prognostic Value of MUC2 Expression in Colorectal Cancer: A Systematic Review and Meta-Analysis. *Gastroenterol Res Pract* (2018) 2018:6986870. doi: 10.1155/2018/6986870
53. Pan X, Wang Q, Xu C, Yan L, Pang S, Gan J. Prognostic value of chloride channel accessory mRNA expression in colon cancer. *Oncol Lett* (2019) 18(3):2967–76. doi: 10.3892/ol.2019.10615
54. Wang L, Gong Y, Saci A, Szabo PM, Martini A, Necchi A, et al. Fibroblast Growth Factor Receptor 3 Alterations and Response to PD-1/PD-L1 Blockade in Patients with Metastatic Urothelial Cancer. *Eur Urol* (2019) 76(5):599–603. doi: 10.1016/j.eururo.2019.06.025
55. Kaprio T, Hagstrom J, Mustonen H, Koskensalo S, Andersson LC, Haglund C. REG4 independently predicts better prognosis in non-mucinous colorectal cancer. *PLoS One* (2014) 9(10):e109600. doi: 10.1371/journal.pone.0109600

**Conflict of Interest:** Authors ZJ, YiZ, YiY, JW and YaY were employed by GloriousMed Clinical Laboratory (Shanghai) Co., Ltd.

The remaining 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.

Copyright © 2021 Jiang, Jin, Zhang, Peng, Zhang, Zhu, Wang, Tong, Yang, Wang, Yang and Xiao. 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.