



# Prognostic value of tumor mutation burden and the relationship between tumor mutation burden and immune infiltration in *HER2+* breast cancer: a gene expression-based study

Ying Wen<sup>1#</sup>, Dengjie Ouyang<sup>1,2#</sup>, Qitong Chen<sup>1</sup>, Liyun Zeng<sup>1</sup>, Na Luo<sup>1</sup>, Hongye He<sup>1</sup>, Munawar Anwar<sup>1</sup>, Limeng Qu<sup>1</sup>, Qiongyan Zou<sup>1</sup>, Wenjun Yi<sup>1</sup>

<sup>1</sup>Department of General Surgery, The Second Xiangya Hospital, Central South University, Changsha, China; <sup>2</sup>Department of General Surgery, Xiangya Hospital, Central South University, Changsha, China

**Contributions:** (I) Conception and design: W Yi, Q Zou; (II) Administrative support: Q Zou; (III) Provision of study materials or patients: Y Wen, D Ouyang; (IV) Collection and assembly of data: Q Chen, L Zeng, N Luo, H He; (V) Data analysis and interpretation: Y Wen, D Ouyang, M Anwar, L Qu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

<sup>#</sup>These authors contributed equally to this work.

**Correspondence to:** Prof. Wenjun Yi; Prof. Qiongyan Zou. Department of General Surgery, The Second Xiangya Hospital, Central South University, Changsha 410011, China. Email: yiwenjun@csu.edu.cn; zqy4311@csu.edu.cn.

**Background:** Whether tumor mutation burden (TMB) correlated with improved survival outcomes or promotion of immunotherapies remained controversy in various malignancies. We aimed to explore the prognostic value of TMB and the relationship between TMB and immune infiltration in *human epidermal growth factor receptor 2*-positive (*HER2+*) breast cancer (BC).

**Methods:** We downloaded somatic mutation data and clinical information for 216 *HER2+* BC patients from the The Cancer Genome Atlas (TCGA) and cBioPortal databases. Patients were divided into high- and low-TMB groups through TMB calculation. Cox regression analysis was used to establish an immune- and mutant-related risk model based on 5-hub genes. The relationship between 5-hub genes mutants and the level of immune infiltration, as well as the relationship between the risk model and the immune microenvironment were analyzed by “TIMER” database.

**Results:** TMB was negatively correlated with overall survival (OS) and disease-free survival (DFS), and high TMB may inhibit immune infiltration in *HER2+* BC. Furthermore, risk score classified effectively patients into low- and high-risk groups in training and validation cohorts. The infiltration of CD4<sup>+</sup> T cells and NK cells and the levels of immune checkpoint pathway genes were lower in the high-risk group, which indicated a poor prognosis.

**Conclusions:** Higher TMB correlated with poor survival outcomes and might inhibit the immune infiltrates in *HER2+* BC. The 5-hub TMB-related signature conferred lower immune cells infiltration which deserved further validation.

**Keywords:** *HER2+* breast cancer (*HER2+* BC); tumor mutation burden (TMB); immune infiltration; prognostic model

Submitted Aug 28, 2021. Accepted for publication Nov 16, 2021.

doi: 10.21037/gs-21-594

**View this article at:** <https://dx.doi.org/10.21037/gs-21-594>

## Introduction

*Human epidermal growth factor receptor-2* (*HER2*) positive breast cancer (BC) accounts for 15–20% of newly diagnosed invasive breast carcinomas (1). At present, anti-*HER2*

therapies, such as trastuzumab and pertuzumab, are the main treatment strategies (2). Moreover, *HER2*-targeting therapeutics have been approved, including drug-antibody conjugated ado-trastuzumab emtansine (T-DM1) and

tyrosine kinase inhibitors (TKIs), such as lapatinib and pyrotinib (3-5). However, due to primary or secondary drug resistance against anti-*HER2* therapies, some *HER2*+ BC patients will eventually progress (6). Thus, there is an urgent need to develop novel therapies and new approaches to overcome the limitations of targeted therapy and improve treatment.

Cancer immunotherapy, such as immune checkpoint blockade (ICB) with nivolumab, pembrolizumab and atezolizumab, has shown increasingly noteworthy value in the treatment of a variety of advanced and refractory cancers (7), such as in the treatment of melanoma (8) and lung cancer (9) and chimeric antigen receptor T cell therapy (ipilimumab) has been successfully utilized to treat refractory leukemia and lymphoma (10). Although BC is classified as a moderately immunogenic cancer (11,12), growing evidence suggests the existence of variable immunogenic activity in BC subtypes (13,14). Safonov *et al.* (15) found that increased immune metagene expression associated significantly with lower clonal heterogeneity in all subtypes of BC and with a trend for lower overall mutation, neoantigen, and CNV loads in TNBC and *HER2*+ cancers. While in estrogen receptor (ER)+ cancers, mutation load, neoantigen load, and CNV load weakly but positively associated with immune infiltration, which reached significance for overall mutation load only. Meanwhile, Schroth *et al.* (16) suggested that BRCA-related DNA repair deficiency and suppressed tumor immune responses may be clinically relevant predictors of endocrine therapy complementing treatment options in subgroups of hormone-sensitive early BC. Additionally, Musolino *et al.* (17) confirmed that the biological characteristics of Fc fragment of IgG receptor IIIa (FCGR3A) can facilitate the efficacy of trastuzumab by enhancing the antibody-dependent cell-mediated cytotoxicity (ADCC) effect. Therefore, based on the promising results from these studies, immunotherapy for *HER2*+ BC is a viable clinical objective.

Nevertheless, these immunotherapies are beneficial to only 20% of cancer patients (18). Thus, many efforts have been devoted to identifying more immunotherapy targets and clarifying molecular mechanism of immunotherapy responsiveness (19). Currently, *programmed cell death-1/programmed cell death-ligand 1 (PD-1/PD-L1)* expression on tumors (20), DNA mismatch-repair deficiency, neoantigen load (21), tumor-infiltrating lymphocytes (22) and IFN- $\gamma$ -related mRNA profile (23) are recognized molecular determinants. However, a preliminary exploratory study has

shown that monotherapy with PD-1/PD-L1 inhibitors is not satisfactory for BC, and the population that benefits from this therapy is limited (24). Tumor mutation burden (TMB) is also a potential biomarker for immunotherapy. Recent studies have found that TMB has a close relationship with immune infiltration and the prognosis of various cancers (25) and the TAPUR study suggested that BC patients in the high-TMB group can benefit from pembrolizumab monotherapy (26). However, the relationship between TMB-related genomic biomarkers and immune infiltration in *HER2*+ BC are still unclear. In this study, we analyzed the somatic mutations and RNA-seq data for *HER2*+ BC in TCGA and divided the samples into high- and low-TMB expression groups. Next, we analyzed the influence of TMB-related differentially expressed genes (DEGs) on the immune microenvironment. Finally, we established a risk model combining immune-related genes and mutant genes to evaluate prognosis for *HER2*+ BC patients. We believe that the risk score system have potential in patient management and that the selected hub genes have the potential to be combined targets for immunotherapy for *HER2*+ BC patients. We present the following article in accordance with the REMARK reporting checklist (available at <https://gs.amegroups.com/article/view/10.21037/ggs-21-594/rc>).

## Methods

### *Acquisition of somatic mutation and RNA-seq data*

We prepared somatic mutation data for 216 *HER2*+ BC from the “Masked Somatic Mutation” category in TCGA database via the GDC data portal (<https://portal.gdc.cancer.gov/>). The “maftools” R package provides multiple analysis modules to perform the visualization process (27). Additionally, the corresponding clinical data sheets for *HER2*+ BC were obtained from cBioPortal (<https://www.cbioportal.org/>) and used as the training dataset. Of 114 *HER2*+ samples from GSE50948 and 323 *HER2*+ samples from GSE96058 in the Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>) were used as the validation datasets. In the GSE50948 dataset, we found that only 63 patients who had received a combination of trastuzumab and chemotherapy (CT) in neoadjuvant therapy, while the other patients received CT alone. Patient inclusion criteria: immunohistochemistry (IHC) or fluorescence in situ hybridization (FISH) was used to determine the status of *HER2* in BC. If one of these was positive, we define these samples as *HER2*+

regardless of ER or progesterone receptor (PR) status. Patient exclusion criteria: (I) both IHC and FISH were negative or uncertain *HER2* status; (II) patients with incomplete clinicopathological information, such as TX stage (the primary tumor could not be assessed), NX stage (regional lymph node involvement could not be assessed), and MX stage (the metastatic status could not be assessed) in the TNM staging system. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

### *Calculation and grouping of TMB*

We defined TMB as the total number of somatic mutations per coding area of a tumor genome or the total number of mutations per megabase (mut/Mb). Perl software was applied to calculate TMB, and only mutations that caused amino acid changes (nonsynonymous mutation) were included. In our cohort, the average TMB (3.05 mut/Mb) was approximately equivalent to the 80th percentile of the TMB distribution. Using the average TMB as the cutoff value, patients were divided into high- (n=43) and low-TMB (n=173) groups. Previous studies had shown that it is feasible to use the average as an initial threshold and verified it as an acceptable threshold choice by alternative cutpoint testing (28).

### *Prognostic analysis*

Kaplan-Meier analyses with log-rank tests were subsequently performed to assess the differential outcomes in overall survival (OS) and disease-free survival (DFS) between the two levels by the “survival” R package. The definition of DFS: in cancer, the length of time after primary treatment for a cancer ends that the patient survives without any signs or symptoms of that cancer. Also called DFS, relapse-free survival, and RFS. In addition, we further evaluated the relationship between the TMB level and multiple clinical variables.

### *ImmuCellAI, TIMER and Immport database*

ImmuCellAI (<http://bioinfo.life.hust.edu.cn/ImmuCellAI#!/>) is a tool that estimates the abundance of 24 types of immune cell infiltrates from a gene expression dataset (29). Based on the “Gene” and “sCNA” modules of the TIMER database (<https://cistrome.shinyapps.io/timer/>), we further evaluated the correlation between hub

genes and immune microenvironment in *HER2+* BCs. Moreover, we obtained 1,811 immune-related genes from the Immport database (<https://www.immport.org/shared/home>).

### *DEGs and functional analysis*

The “edgeR” R package was utilized to identify DEGs between the two groups with  $|\log(\text{fold change})| > 1$  and  $\text{FDR} < 0.05$ . The “Venn” diagram identified 29 immune-related DEGs. We performed Gene Ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis on DEGs using the David database (<https://david.ncicrf.gov/>). The difference between high- and low-TMB cohorts can be detected by the “maftools” R package, which performs a Fisher test on all genes to detect differentially mutant genes. We selected the first 30 mutant genes with significant differences.

### *Construction and validation of the immune- and mutant-related risk scoring system*

After excluding patients with missing mutation information and distant metastasis, 178 samples were subjected to subsequent analysis. The clinical characteristics of all patients are shown in *Table 1*. In the immune- and mutant-related genes identified above, we finally obtained 17 immune genes and 6 mutant genes with prognostic significance in the Kaplan-Meier plotter database (<https://kmplot.com/analysis/>) (30). We used the Cox proportional hazard regression analysis (Cox regression analysis) to construct a risk scoring system and multiplied by the gene expression of each patient. A receiver operating characteristic (ROC) curve was drawn to assess the predictive power of the system using the “survivalROC” R package, and a Kaplan-Meier analysis with log-rank test was performed using the “survminer” R package.

### *Statistical analysis*

The “Limma” package was utilized to conduct the normalization and differential analysis. The Wilcoxon rank-sum test was a nonparametric statistical test mainly utilized for comparing two groups, and the Kruskal-Wallis test was suitable for comparing two or more groups. The Chi-square test was used to compare composition ratios or rates between the two groups. All statistical analysis was implemented based on R software (Version 4.0.2, <https://>

**Table 1** Clinical baseline of all 178 *HER2+* BC patients from TCGA cohort

Variables	TCGA cohort (n=178)	
	n	%
Status	153	85.96%
Alive	153	85.96%
Dead	25	14.04%
Age, years		
≥60	81	45.51%
<60	97	54.49%
AJCC-T		
T1	36	20.22%
T2	117	65.73%
T3	15	8.43%
T4	10	5.62%
AJCC-N		
N0	74	41.57%
N1	68	38.20%
N2	25	14.04%
N3	11	6.18%
AJCC-M		
M0	174	97.75%
M1	4	2.25%
Stage		
Stage I	20	11.24%
Stage II	109	61.24%
Stage III	45	25.28%
Stage IV	4	2.25%
ER status		
Positive	123	69.10%
Negative	55	30.90%
PR status		
Positive	95	53.37%
Negative	83	46.63%

BC, breast cancer; TCGA, The Cancer Genome Atlas; ER, estrogen receptor; PR, progesterone receptor.

www.r-project.org/). The conventional two-sided P value <0.05 was considered significant.

## Results

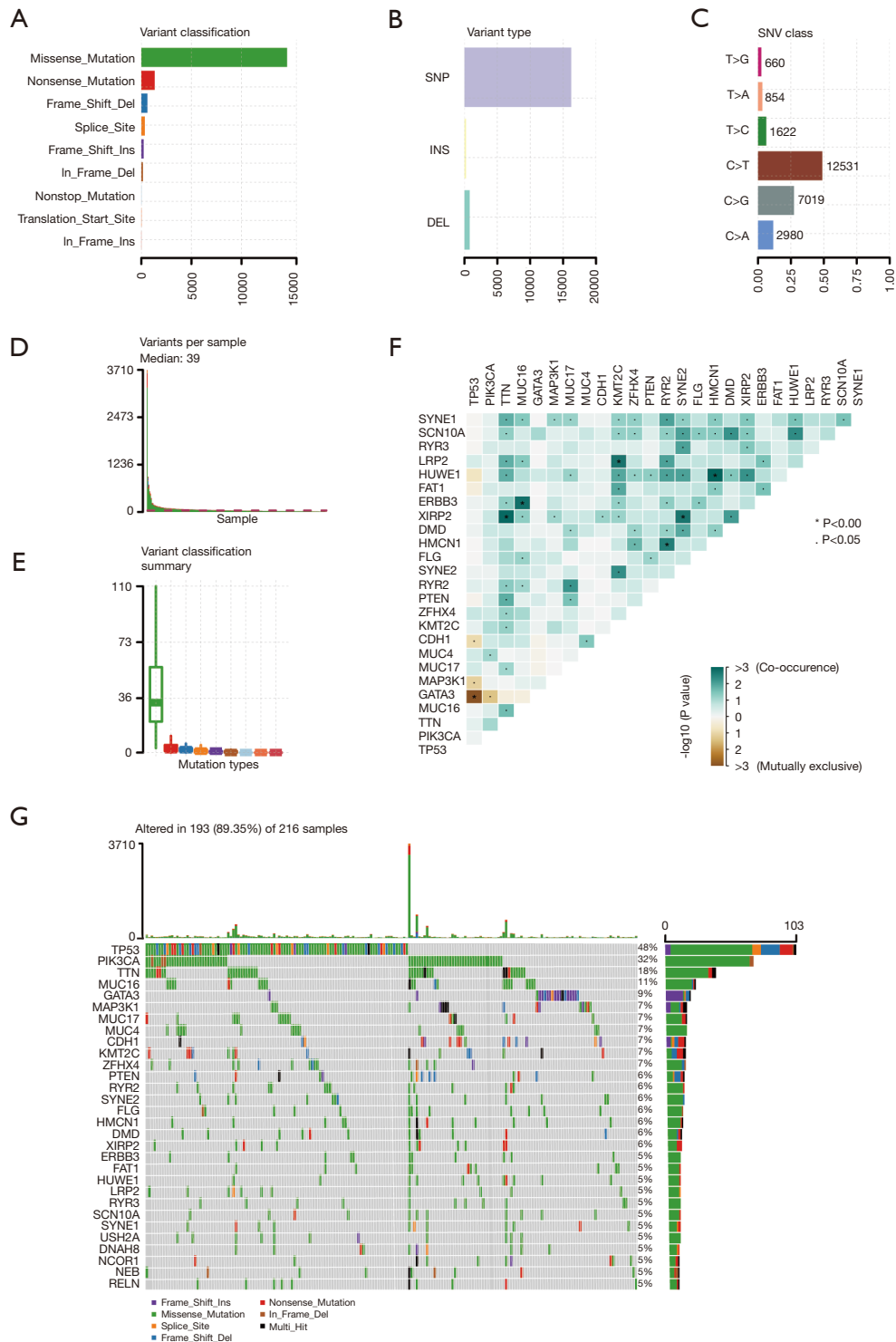
### *Landscape of the HER2+ BC mutation profiles*

In summary, various mutation categories were categorized into different groups, in which missense mutations were the most common type, single nucleotide polymorphisms (SNPs) occurred more frequently than insertions or deletions, and C > T was the most common single nucleotide variant (SNV) (*Figure 1A-1C*). Moreover, we calculated the number of mutant bases in each sample and displayed the mutation categories using different colors in box plots (*Figure 1D,1E*). The coincident and exclusive associations were shown across mutant genes (*Figure 1F*). Mutation information for each gene in each sample is shown in a waterfall plot, and *tumor protein P53 (TP53; 48%)*, *phosphoinositide 3-kinase alpha (PIK3CA; 32%)* and *titin (TTN; 18%)* were the top three mutant genes in the samples (*Figure 1G*).

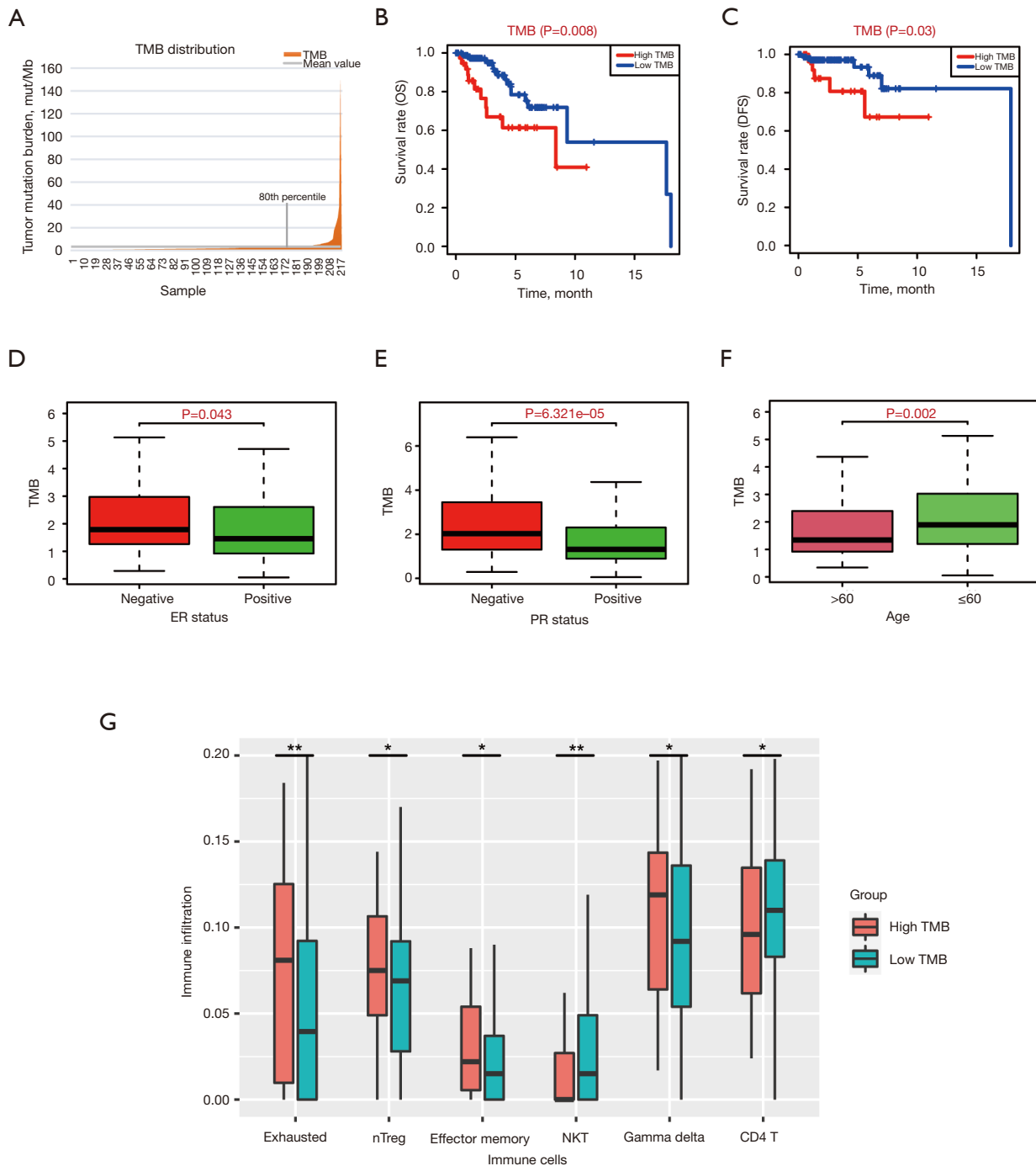
### *TMB is correlated with survival, hormone receptor and immune infiltration of HER2+ BC*

Using the average TMB as the cutoff point, we divided all samples into high- and low-TMB groups (*Figure 2A*). A higher level of TMB correlated with worse OS (P=0.008) and conferred a higher risk of tumor recurrence (P=0.03) (*Figure 2B,2C*). Moreover, a high TMB level was significantly related to hormone receptor negative status and advanced age (*Figure 2D-2F*). Nevertheless, no significant difference was found in TNM stages or pathological stages (*Figure S1*).

Based on the ImmuCellAI database, we estimated the abundance of 24 immune cells in each patient. In the high-TMB group, the infiltration levels of exhausted T cells, nature T regulatory cells (nTreg), effector memory T cells and gamma delta T cells were higher than that of the low-TMB group, while CD4<sup>+</sup> T cells and Natural Killer (NK) cells were the opposite (*Figure 2G*). However, no significant differences were observed in the infiltration levels of other immune cells between the two TMB groups (*Figure S2*).



**Figure 1** Landscape of the *HER2+* BC mutation profiles. (A-C) Classification of mutation types according to different categories, in which missense mutation accounts for the largest fraction, SNP had a higher frequency than insertion or deletion, and C > T was the most common SNV. (D,E) Tumor mutation burden in each sample. (F) The coincident and exclusive associations across mutant genes. (G) Mutation information of each gene in each sample was shown in the waterfall plot. BC, breast cancer; SNP, single nucleotide polymorphism; SNV, single nucleotide variant.



**Figure 2** TMB correlated with survival, hormone receptor and immune infiltration of *HER2+* BC. (A) The TMB distribution for the TCGA cohort was shown. The mean TMB (3.05 mut/Mb) was marked by a gray line and coincides with the 80th percentile. (B,C) Higher TMB levels correlated with poor OS and DFS. (D-F) Higher TMB level was associated with hormone receptor negativity and advanced age. (G) TMB correlated with immunocyte infiltration of *HER2+* BC. \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ . TMB, tumor mutation burden; OS, overall survival; DFS, disease-free survival; BC, breast cancer; TCGA, The Cancer Genome Atlas.



### Identification of immune- and mutant-related DEGs between high- and low-TMB groups

A total of 182 DEGs (Table S1) were identified and are shown in a volcano plot (Figure 3A). We screened 29 immune-related DEGs in 1,811 immune genes from the Immport database by the “Venn” diagram (Figure 3B) and further identified 17 genes with prognostic significance in the Kaplan-Meier plotter database (Figure S3). In the biological process group, some immune processes involved in the functional analysis were the response to cytokines and the negative regulation of leukocyte differentiation (Figure 3C). KEGG pathway analysis indicated that these DEGs were mainly involved in the Rap1 signaling pathway (Figure 3D).

“Co-oncoplot” was used to map the two TMB groups side by side to better compare the differentially mutant genes, with the frequency and type of mutations shown (Figure 3E). We selected the first 30 of the 91 differentially mutant genes (Table S2) for prognostic analysis of mutation and expression, and finally, 6 mutant genes were identified (Figure S4).

### Construction and validation of the immune- and mutant-related risk scoring system

After identifying 17 immune- and 6 mutant-related genes, we used the Cox regression analysis to construct an immune- and mutant-related risk scoring system including 5 hub genes. We performed univariate Cox regression analysis on each of the characteristics and for factors with a P value <0.10 (Table S3), then we further constructed the final risk model by multivariate Cox regression analysis. The risk scores formula is as follows: risk score =  $0.860788 \times \exp^{DNAH11} + 0.505288 \times \exp^{ERBB3} - 0.41573 \times \exp^{FABP5} - 0.43514 \times \exp^{MAPT} - 0.60184 \times \exp^{PGR}$  (Table 2). Then, the risk score of each sample was calculated. Patients in the high-risk group had worse survival outcomes ( $P=2.205e-04$ ) (Figure 4A). The ROC curve of 5-year OS prediction was drawn to assess the predictive accuracy with area under curve (AUC) =0.759 (Figure 4B). Moreover, we also use 17 immune-related genes and 6 mutant-related genes to establish risk models, respectively (Figure S5A,S5B, Tables S4,S5). Altogether, the combined model was indeed better than the above two alone.

We utilized the univariate and multivariate Cox analysis to investigate the relationship of the risk score and clinical factors. Multivariate Cox analysis suggested that age and

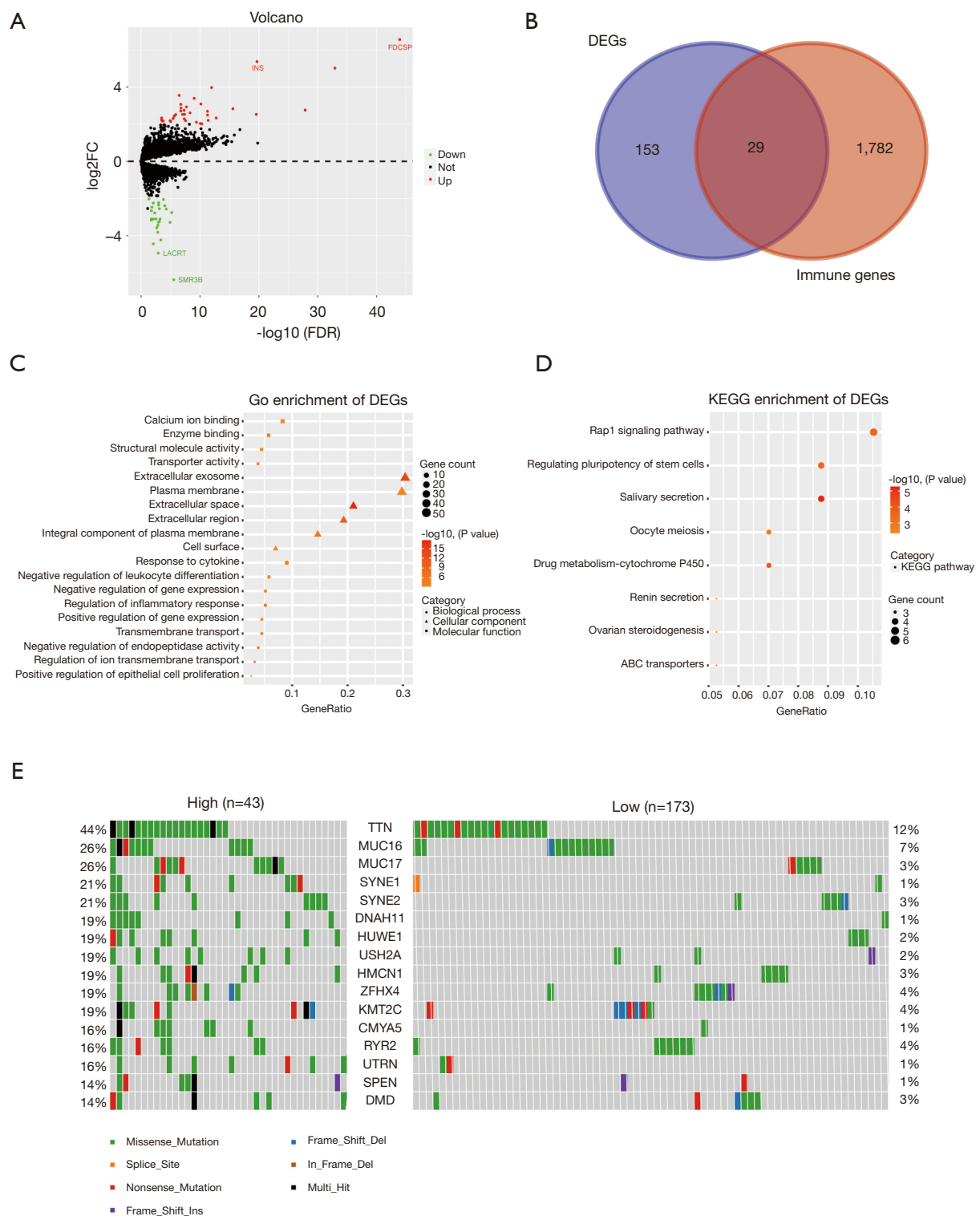
risk score are independent prognostic factors, but TMB was not (Figure 4C,4D). In addition, higher risk was related to higher TMB and lower *PD-1*, *PD-L1*, and *CTLA-4* (Figure 4E). After grouping by risk value in the GSE50948 dataset, we found only 63 patients who had received a combination of trastuzumab and CT in neoadjuvant therapy. The pathologic complete response (pCR) of the high-risk group was 61.29%, while that of low-risk group was 38.7%. Moreover, the pCR was 53.85% in the high-risk group and 46.15% in the low-risk group for those who received CT alone (Figure 4F). The heat map shows the expression of the 5 hub genes and clinicopathological variables in the high- and low-risk groups (Figure 5). We verified that the expression level of *ERBB3*, *DNAH11* in high-risk group was higher than low-risk group, while *MAPT*, *FABP5*, *PGR* were the opposite.

To verify that the system has good prognostic prediction capabilities, we screened out 323 *HER2+* BC from the GSE96058 dataset and then applied the same risk scoring formula. High-risk patients had significantly worse prognosis ( $P=2.417e-02$ ). ROC analysis showed that the system has a relatively good prognostic prediction (AUC =0.609) (Figure S5C).

### Immune infiltration between the low- and high-risk groups

We explored the potential relationship between our risk scoring system and immune infiltration using the “ImmuCellAI” tool. Most of the immune cells in the high-risk group had a lower degree of infiltration, including  $CD4^+$  T cells,  $CD8^+$  T cells, NK cells, exhausted T cells, T regulatory type 1 cells (Tr1), follicular helper T cells (Tfh), mucosal-associated invariant T cells (MAIT), gamma delta T cells and cytotoxic T cells (Tc), in which  $CD4^+$  T and NK cells were consistent with the previous results of immune infiltration between high- and low-TMB groups. However, the higher infiltration degree of monocyte, neutrophil and helper T 17 cells (Th17) were observed in the high-risk group (Figure 6).

Based on the TIMER database, we further found that various forms of mutations carried by hub genes can usually suppress immune infiltration, including  $CD4^+$  and  $CD8^+$  T cells, neutrophils, macrophages, dendritic cells and NK cells (Figure S6). However, the expression of *DNAH11* was negatively correlated with the infiltration level of immune cells, and the expression of other genes (except *ERBB3*) was positively correlated (Figure S7).



**Figure 3** Comparison of gene profiles in high- and low- TMB groups and enrichment analysis. (A) The volcano plot of DEGs. (B) Identification of 29 TMB-related immune genes. (C,D) GO and KEGG results revealed that DEGs were involved in some immune-related processes. (E) The “Co-oncoplot” of differentially mutant genes in two TMB groups. DEGs, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; TMB, tumor mutation burden.



**Table 2** The immune- and mutant- related risk scoring system (the combined model)

Genes	Coef	HR	HR.95L	HR.95H	P value
<i>DNAH11</i>	0.860788	2.365025	1.191053	4.696132	0.013912
<i>ERBB3</i>	0.505288	1.657463	0.992778	2.767167	0.053329
<i>FABP5</i>	-0.41573	0.659858	0.379027	1.148763	0.141647
<i>MAPT</i>	-0.43514	0.647174	0.45371	0.923132	0.016334
<i>PGR</i>	-0.60184	0.547805	0.258784	1.159615	0.115735

## Discussion

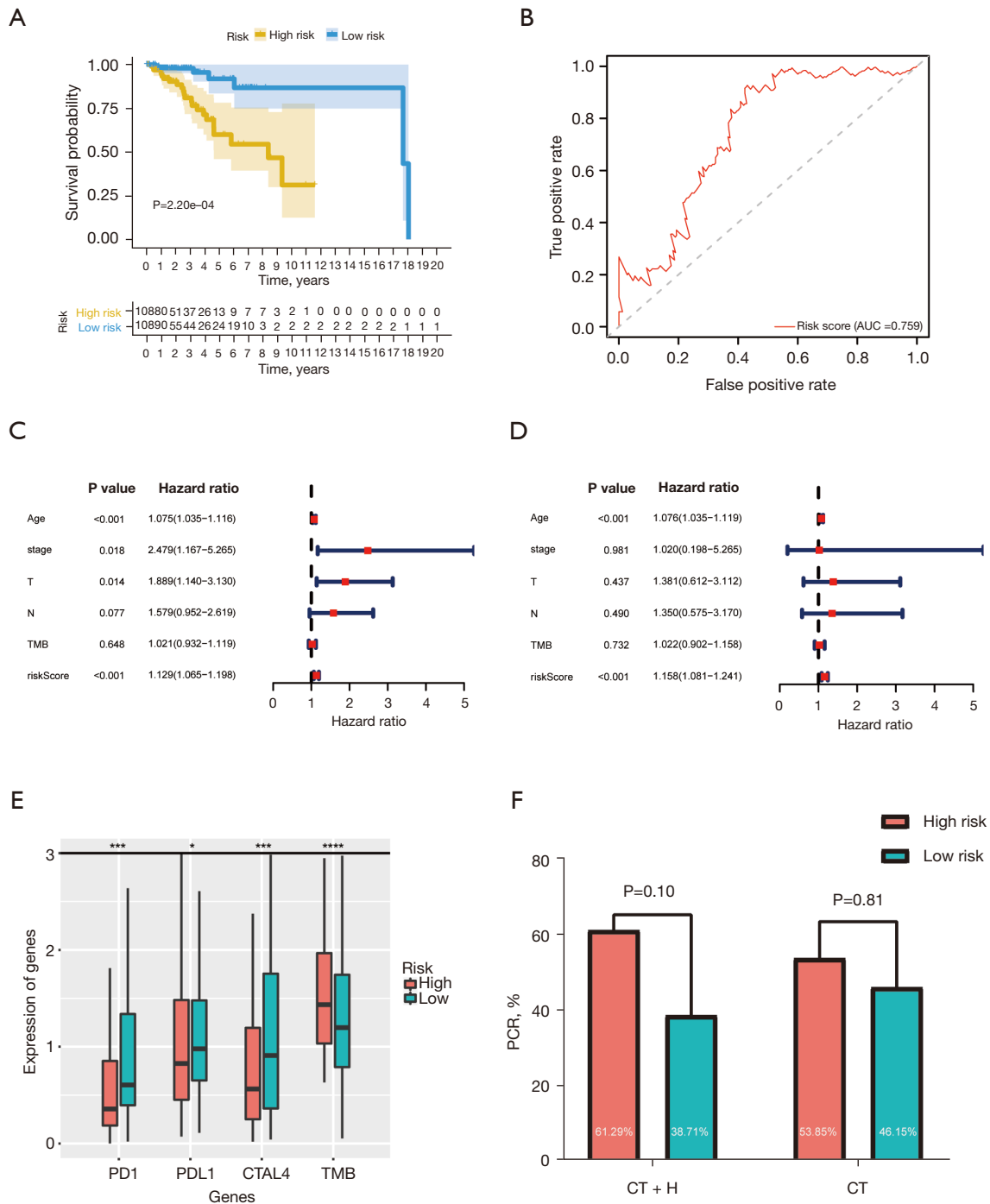
In this study, we comprehensively studied the role of TMB in the regulation of the immune phenotype in *HER2+* BC. Then, a risk scoring system related to immunity and mutation was developed based on the TCGA dataset. This risk scoring system has favorable prognostic prediction ability, which is independent of traditional prognostic factors. Furthermore, we explored the correlation between the risk score and immune cell infiltration in *HER2+* BC. To our knowledge, this is the first prognostic model combining immune genes with mutant genes to predict survival outcomes in *HER2+* BC.

TMB has been described as a predictor of tumor behavior and immunological response (31-33), and it has been demonstrated to be effective in a variety of tumors, such as colorectal cancer (34), lung cancer (35,36), etc. In our study, higher TMB was significantly correlated with worse OS and DFS. However, a recent bioinformatics study demonstrated that in *HER2+* refractory metastatic BC, patients with lower TMB had worse survival outcome than those who had higher TMB (44.9 vs. 85.8 months,  $P=0.016$ ) (37). The difference between the two studies is that the patients of the two studies are two independent cohorts. Samples for our study were derived from TCGA dataset, and patients at advanced stages were excluded. In addition, current studies have shown that trastuzumab can change the immune status of *HER2+* BC by enhancing the ADCC effect, making patients more likely to benefit from this treatment (17,38). Only a small number of patients in our cohort received trastuzumab treatment. Thus, the treatment may lead to different levels of TMB expression and different prognosis. Our results indicated that TMB is a potential biomarker of immune-mediated survival in *HER2+* BC.

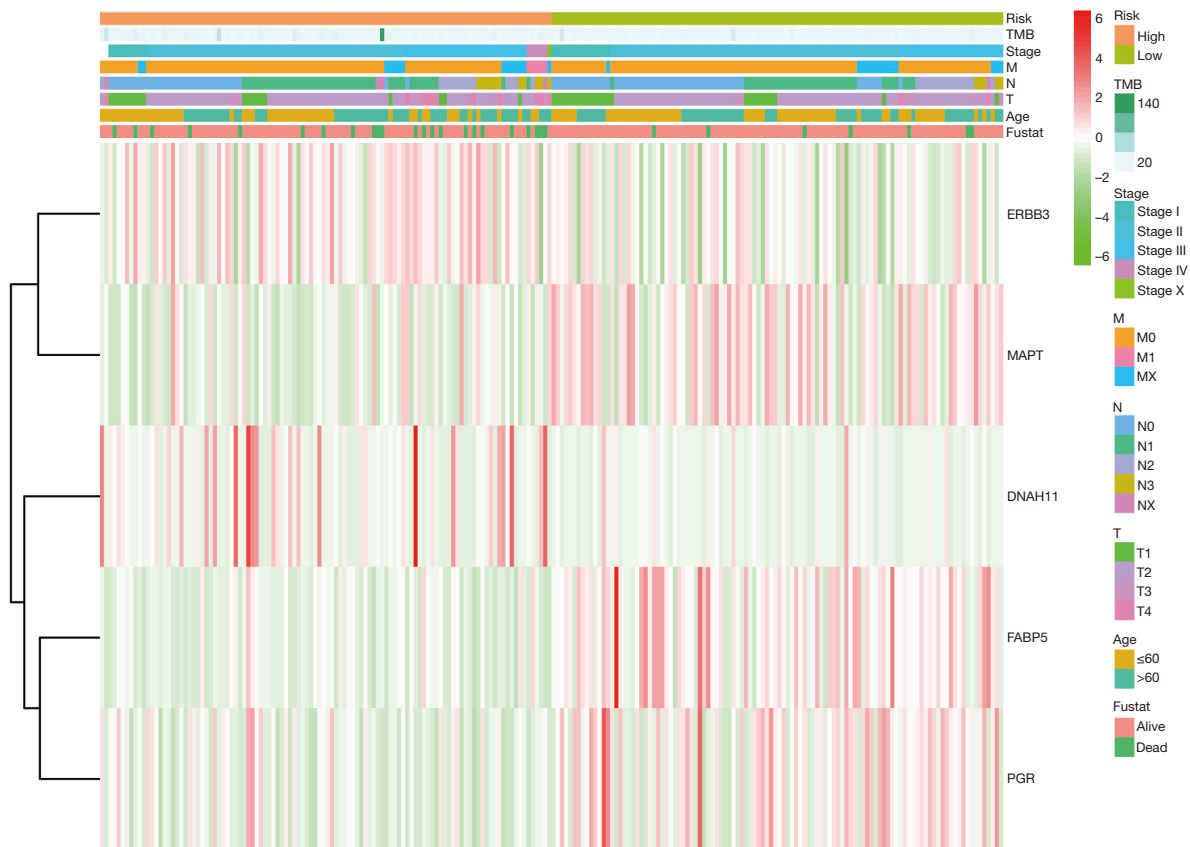
Then, we identified five hub genes, including two mutant genes (*DNAH11*, *ERBB3*) and three immune-related genes (*FABP5*, *MAPT*, *PGR*), that were highly associated with OS in *HER2+* BC. Then, a predictive model was developed.

Patients with a high-risk score had poor survival outcomes, and the AUC of the ROC curve was 0.759, indicating the superior predictive accuracy. In our study, there were no mutations in differentially expressed immune-related genes. Therefore, we used Cox regression analysis to perform immune-related gene modeling (AUC =0.741) and mutant gene modeling (AUC =0.624). The result indicated that the prognosis modeled by only one class of genes was not as good as combination modeling. In short, our research proposed a combination modeling to predict the prognosis in *HER2+* BC and requires validation in the future.

Furthermore, we analyzed the association between the risk score and the expression of critical immune checkpoint genes. The results showed that patients in the low-risk group had higher *PD-1*, *PD-L1* and *CTLA-4* expression. Thus, the patients in the low-risk group were more likely to benefit from ICBs. However, these results did not match with the low-TMB expression in the low-risk group. At present, the expression level of *PD-L1* can be evaluated by immunohistochemistry (IHC), but there are still limitations to using *PD-L1* as a biomarker of immunotherapy response. *PD-L1* expression in tumors is heterogeneous and is affected by previous treatments, such as CT and radiotherapy (39,40). Furthermore, in the KEYNOTE-055 trial, researchers found that *PD-L1*-negative patients also had a significant response to pembrolizumab treatment, suggesting that evaluation of the immunotherapy response with *PD-L1* alone is not sufficient (41). Importantly, some studies have also demonstrated that TMB has no association with *PD-L1* expression, indicating that TMB and *PD-L1* expression were independent biomarkers for predicting the response to ICBs (42,43). Some studies now showed that trastuzumab in combination with ICBs may benefit some patients. Su *et al.* (44) confirmed that trastuzumab can enhance the expression of *PD-L1*. The PANACEA study estimated that in the *PD-L1*-positive subgroup, the objective response



**Figure 4** Construction and assessment of the combined risk scoring system. (A) Kaplan-Meier analysis demonstrated that patients with a higher risk score had worse OS. (B) ROC curve of the combined model in the training cohort. (C,D) The univariate and multivariate Cox analysis in *HER2+* BC. The hazard ratio with 95% CI for each independent TMB-related signature was shown in a forest plot. (E) TMB was significantly correlated with immune checkpoint. (F) Comparison of pCR rate in high- and low-risk patients after neoadjuvant CT in the NOAH clinical trial (GSE50948). \*, P<0.05; \*\*\*, P<0.001; \*\*\*\*, P<0.0001. AUC, area under curve; TMB, tumor mutation burden; pCR, pathologic complete response; CT + H, chemotherapy plus Herceptin; CT, chemotherapy; OS, overall survival; ROC, receiver operating characteristic; BC, breast cancer; CI, confidence interval.



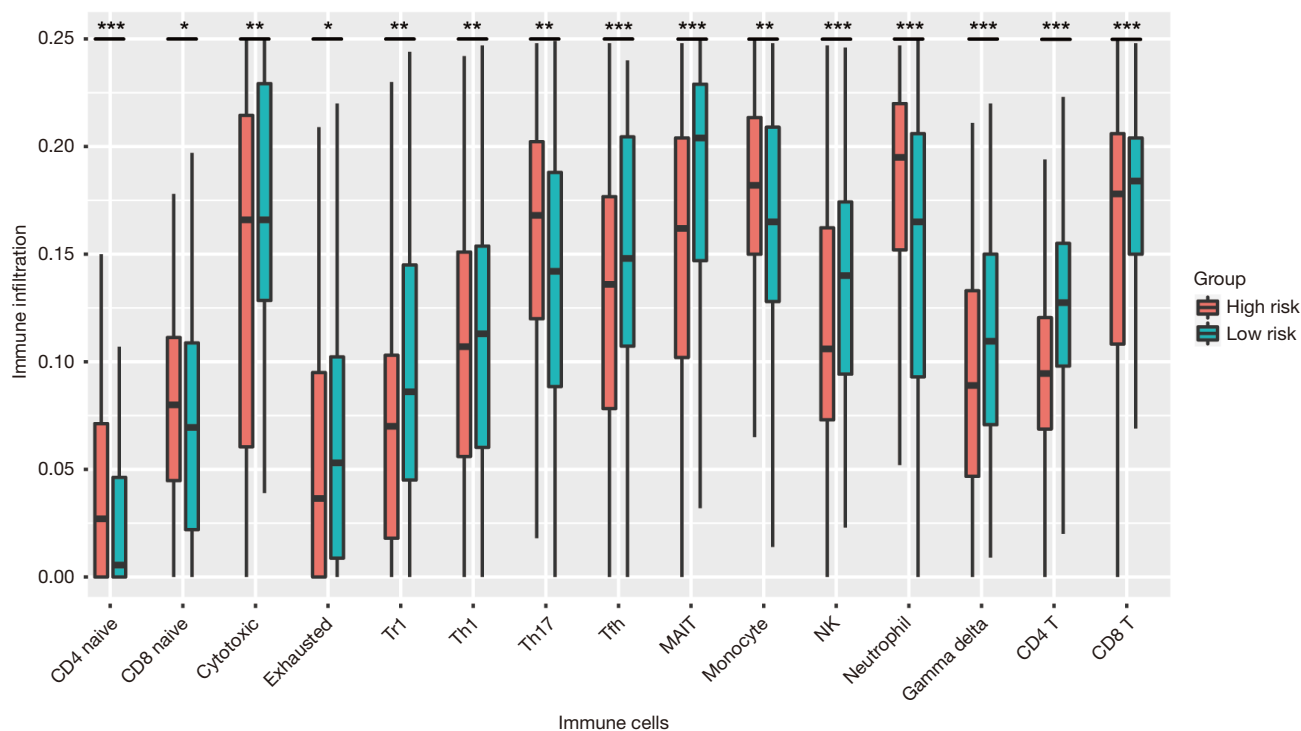
**Figure 5** The heat map shows the expression of 5 signatures and the distribution of clinicopathological variables between the high- and low-risk groups. TMB, tumor mutation burden.

rate (ORR) of trastuzumab combined with pembrolizumab was 15% (45); similarly, the KATE2 study showed that the 1-year OS rate of patients in the atezolizumab group was better (46). However, our study suggested that patients in the high-risk group were able to achieve a higher pCR trend with trastuzumab plus CT, and this is inconsistent with poorer outcomes in the high-risk group. Our results may indicate that patients in the high-risk group are more sensitive to CT plus trastuzumab, which may be not related to the treatment of trastuzumab, but be associated with the change of immune microenvironment. pCR rate is not the only prognostic factors, as clinical trial of NeoALTTO, for example, suggested no difference in OS and DFS between patients in the pCR group and those in the non-PCR group (47). Thus, more studies are investigating changes in immune markers after trastuzumab treatment to explore more methods and strategies for combined therapy. In terms of immune cell infiltration, more regulatory T cells and fewer NK cells were found in the

high-TMB/high-risk group, consistent with similar results in another study of *HER2+* BC, where more Treg cells and fewer NK cells demonstrated a worse prognosis (48). In summary, our study systematically explored the relationships between TMB and immune infiltration and constructed a TMB-related risk scoring system in *HER2+* BC. Thus, our study provides new insights into the immune cell infiltration of the tumor microenvironment and immunotherapies for *HER2+* BC.

There were some limitations to our study. First, this is a retrospective study, and the results should be further confirmed by prospective studies to exclude sample selection bias and inter-tumor heterogeneity. Additionally, the clinical relevance of the 5 hub genes need more independent validation in better-documented clinical trials in the future.

In conclusion, *HER2+* BC patients with low TMB have a better prognosis than patients with high TMB. TMB could regulate immune cell infiltration. The immune- and mutant-related risk scoring system can divide patients into



**Figure 6** Immune infiltration landscape in the low- and high-risk groups. \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ .

low- and high-risk groups with different outcomes and immunophenotypes.

## Conclusions

Our study suggested that TMB is associated with poor prognosis of *HER2* positive BC, and established a prognostic model for the combination of mutation-related genes and immune genes. Moreover, our results suggested that high-risk group are more likely to benefit from ICIs treatment, which provide a new strategy for evaluating the efficacy of immunotherapy and a new insight for finding *HER2* positive BC patients who are not sensitive to immunotherapy.

## Acknowledgments

**Funding:** This work was supported by grants from the Natural Science Foundation of the Hunan Province of China (2020JJ4828) and the Science and Technology Innovation Program of Hunan Province (2021SK2026).

## Footnote

**Reporting Checklist:** The authors have completed the REMARK reporting checklist. Available at <https://gs.amegroups.com/article/view/10.21037/gS-21-594/rc>

**Peer Review File:** Available at <https://gs.amegroups.com/article/view/10.21037/gS-21-594/prf>

**Conflicts of Interest:** All authors have completed the ICMJE uniform disclosure form (available at <https://gs.amegroups.com/article/view/10.21037/gS-21-594/coif>). All authors report that this work was supported by the Natural Science Foundation of the Hunan Province of China (2020JJ4828) and the Science and Technology Innovation Program of Hunan Province (2021SK2026). The authors have no other conflicts of interest to declare.

**Ethical Statement:** The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are

appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

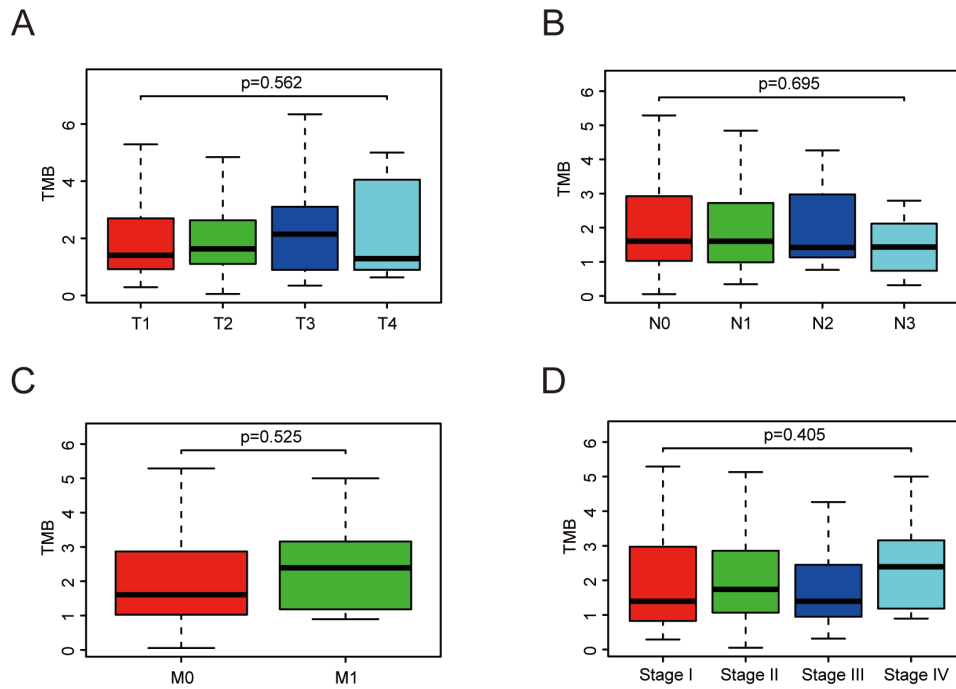
- Chen X, Yuan Y, Gu Z, et al. Accuracy of estrogen receptor, progesterone receptor, and HER2 status between core needle and open excision biopsy in breast cancer: a meta-analysis. *Breast Cancer Res Treat* 2012;134:957-67.
- Swain SM, Kim SB, Cortés J, et al. Pertuzumab, trastuzumab, and docetaxel for HER2-positive metastatic breast cancer (CLEOPATRA study): overall survival results from a randomised, double-blind, placebo-controlled, phase 3 study. *Lancet Oncol* 2013;14:461-71.
- Chen Q, Ouyang D, Anwar M, et al. Effectiveness and Safety of Pyrotinib, and Association of Biomarker With Progression-Free Survival in Patients With HER2-Positive Metastatic Breast Cancer: A Real-World, Multicentre Analysis. *Front Oncol* 2020;10:811.
- Ouyang DJ, Chen QT, Anwar M, et al. The Efficacy of Pyrotinib as a Third- or Higher-Line Treatment in HER2-Positive Metastatic Breast Cancer Patients Exposed to Lapatinib Compared to Lapatinib-Naive Patients: A Real-World Study. *Front Pharmacol* 2021;12:682568.
- Anwar M, Chen Q, Ouyang D, et al. Pyrotinib Treatment in Patients With HER2-positive Metastatic Breast Cancer and Brain Metastasis: Exploratory Final Analysis of Real-World, Multicenter Data. *Clin Cancer Res* 2021;27:4634-41.
- Chung A, Cui X, Audeh W, et al. Current status of anti-human epidermal growth factor receptor 2 therapies: predicting and overcoming herceptin resistance. *Clin Breast Cancer* 2013;13:223-32.
- Siegel RL, Miller KD, Jemal A. *Cancer Statistics, 2017*. *CA Cancer J Clin* 2017;67:7-30.
- Pollack MH, Betof A, Dearden H, et al. Safety of resuming anti-PD-1 in patients with immune-related adverse events (irAEs) during combined anti-CTLA-4 and anti-PD1 in metastatic melanoma. *Ann Oncol* 2018;29:250-5.
- Gong X, Li X, Jiang T, et al. Combined Radiotherapy and Anti-PD-L1 Antibody Synergistically Enhances Antitumor Effect in Non-Small Cell Lung Cancer. *J Thorac Oncol* 2017;12:1085-97.
- June CH, O'Connor RS, Kawalekar OU, et al. CAR T cell immunotherapy for human cancer. *Science* 2018;359:1361-5.
- Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature* 2013;500:415-21.
- Vogelstein B, Papadopoulos N, Velculescu VE, et al. Cancer genome landscapes. *Science* 2013;339:1546-58.
- Lehmann BD, Bauer JA, Chen X, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest* 2011;121:2750-67.
- Miller LD, Chou JA, Black MA, et al. Immunogenic Subtypes of Breast Cancer Delineated by Gene Classifiers of Immune Responsiveness. *Cancer Immunol Res* 2016;4:600-10.
- Safonov A, Jiang T, Bianchini G, et al. Immune Gene Expression Is Associated with Genomic Aberrations in Breast Cancer. *Cancer Res* 2017;77:3317-24.
- Schroth W, Büttner FA, Kandabarau S, et al. Gene Expression Signatures of BRCAness and Tumor Inflammation Define Subgroups of Early-Stage Hormone Receptor-Positive Breast Cancer Patients. *Clin Cancer Res* 2020;26:6523-34.
- Musulino A, Naldi N, Bortesi B, et al. Immunoglobulin G fragment C receptor polymorphisms and clinical efficacy of trastuzumab-based therapy in patients with HER2/neu-positive metastatic breast cancer. *J Clin Oncol* 2008;26:1789-96.
- Braun DA, Burke KP, Van Allen EM. Genomic Approaches to Understanding Response and Resistance to Immunotherapy. *Clin Cancer Res* 2016;22:5642-50.
- Gibney GT, Weiner LM, Atkins MB. Predictive biomarkers for checkpoint inhibitor-based immunotherapy. *Lancet Oncol* 2016;17:e542-51.
- Patel SP, Kurzrock R. PD-L1 Expression as a Predictive Biomarker in Cancer Immunotherapy. *Mol Cancer Ther* 2015;14:847-56.
- Efremova M, Finotello F, Rieder D, et al. Neoantigens Generated by Individual Mutations and Their Role in Cancer Immunity and Immunotherapy. *Front Immunol* 2017;8:1679.



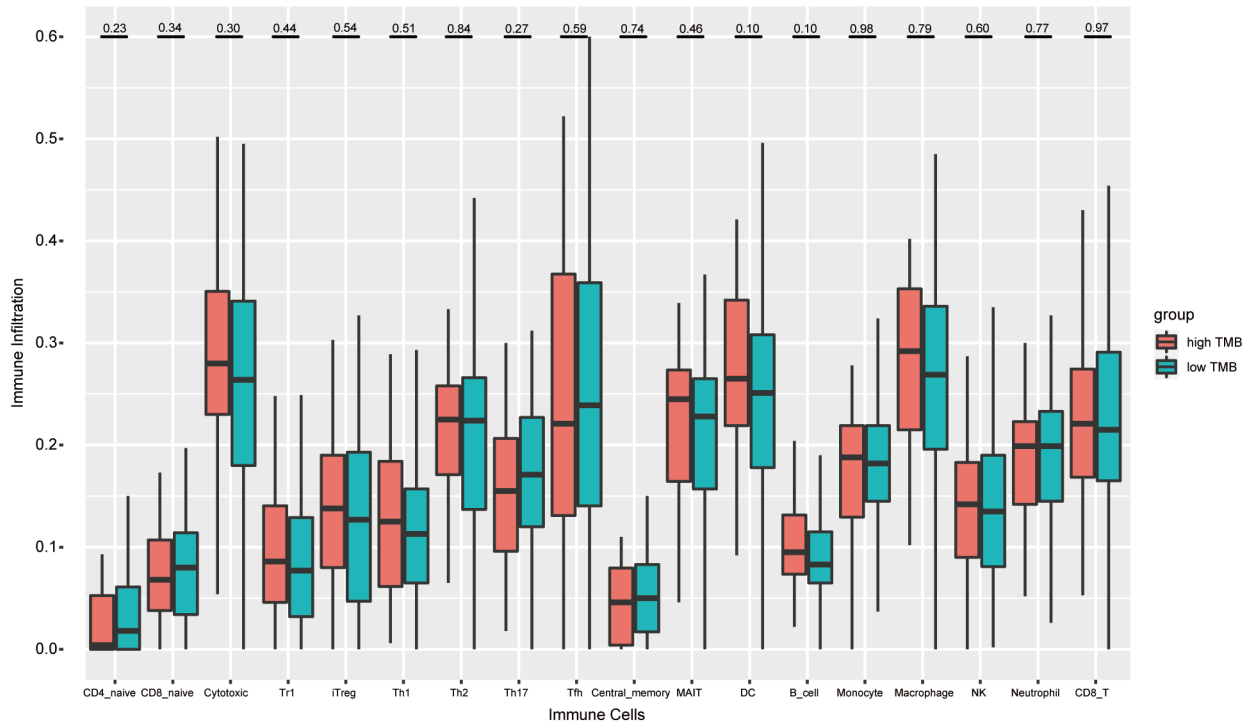
22. Zito Marino F, Ascierto PA, Rossi G, et al. Are tumor-infiltrating lymphocytes protagonists or background actors in patient selection for cancer immunotherapy? *Expert Opin Biol Ther* 2017;17:735-46.
23. Ayers M, Lunceford J, Nebozhyn M, et al. IFN- $\gamma$ -related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest* 2017;127:2930-40.
24. Nanda R, Chow LQ, Dees EC, et al. Pembrolizumab in Patients With Advanced Triple-Negative Breast Cancer: Phase Ib KEYNOTE-012 Study. *J Clin Oncol* 2016;34:2460-7.
25. Wang X, Li M. Correlate tumor mutation burden with immune signatures in human cancers. *BMC Immunol* 2019;20:4.
26. Alva AS, Mangat PK, Garrett-Mayer E, et al. Pembrolizumab in Patients With Metastatic Breast Cancer With High Tumor Mutational Burden: Results From the Targeted Agent and Profiling Utilization Registry (TAPUR) Study. *J Clin Oncol* 2021;39:2443-51.
27. Mayakonda A, Lin DC, Assenov Y, et al. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* 2018;28:1747-56.
28. Thomas A, Routh ED, Pullikuth A, et al. Tumor mutational burden is a determinant of immune-mediated survival in breast cancer. *Oncoimmunology* 2018;7:e1490854.
29. Miao YR, Zhang Q, Lei Q, et al. ImmuCellAI: A Unique Method for Comprehensive T-Cell Subsets Abundance Prediction and its Application in Cancer Immunotherapy. *Adv Sci (Weinh)* 2020;7:1902880.
30. Lanczky A, Gyorffy B. Web-Based Survival Analysis Tool Tailored for Medical Research (KMplot): Development and Implementation. *J Med Internet Res* 2021;23:e27633.
31. Goodman AM, Kato S, Bazhenova L, et al. Tumor Mutational Burden as an Independent Predictor of Response to Immunotherapy in Diverse Cancers. *Mol Cancer Ther* 2017;16:2598-608.
32. Birkbak NJ, Kochupurakkal B, Izarzugaza JM, et al. Tumor mutation burden forecasts outcome in ovarian cancer with BRCA1 or BRCA2 mutations. *PLoS One* 2013;8:e80023.
33. Rizvi NA, Hellmann MD, Snyder A, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 2015;348:124-8.
34. Schrock AB, Ouyang C, Sandhu J, et al. Tumor mutational burden is predictive of response to immune checkpoint inhibitors in MSI-high metastatic colorectal cancer. *Ann Oncol* 2019;30:1096-103.
35. Hellmann MD, Callahan MK, Awad MM, et al. Tumor Mutational Burden and Efficacy of Nivolumab Monotherapy and in Combination with Ipilimumab in Small-Cell Lung Cancer. *Cancer Cell* 2018;33:853-861.e4.
36. Hellmann MD, Nathanson T, Rizvi H, et al. Genomic Features of Response to Combination Immunotherapy in Patients with Advanced Non-Small-Cell Lung Cancer. *Cancer Cell* 2018;33:843-852.e4.
37. Park SE, Park K, Lee E, et al. Clinical implication of tumor mutational burden in patients with HER2-positive refractory metastatic breast cancer. *Oncoimmunology* 2018;7:e1466768.
38. Catenacci DVT, Kang YK, Park H, et al. Margetuximab plus pembrolizumab in patients with previously treated, HER2-positive gastro-oesophageal adenocarcinoma (CP-MGAH22-05): a single-arm, phase 1b-2 trial. *Lancet Oncol* 2020;21:1066-76.
39. Mansfield AS, Aubry MC, Moser JC, et al. Temporal and spatial discordance of programmed cell death-ligand 1 expression and lymphocyte tumor infiltration between paired primary lesions and brain metastases in lung cancer. *Ann Oncol* 2016;27:1953-8.
40. Peng J, Hamanishi J, Matsumura N, et al. Chemotherapy Induces Programmed Cell Death-Ligand 1 Overexpression via the Nuclear Factor- $\kappa$ B to Foster an Immunosuppressive Tumor Microenvironment in Ovarian Cancer. *Cancer Res* 2015;75:5034-45.
41. Bauml J, Seiwert TY, Pfister DG, et al. Pembrolizumab for Platinum- and Cetuximab-Refractory Head and Neck Cancer: Results From a Single-Arm, Phase II Study. *J Clin Oncol* 2017;35:1542-9.
42. Seiwert TY, Burtneß B, Mehra R, et al. Safety and clinical activity of pembrolizumab for treatment of recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-012): an open-label, multicentre, phase 1b trial. *Lancet Oncol* 2016;17:956-65.
43. Cristescu R, Mogg R, Ayers M, et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science* 2018;362:eaar3593.
44. Su S, Zhao J, Xing Y, et al. Immune Checkpoint Inhibition Overcomes ADCP-Induced Immunosuppression by Macrophages. *Cell* 2018;175:442-457.e23.
45. Loi S, Giobbie-Hurder A, Gombos A, et al. Pembrolizumab plus trastuzumab in trastuzumab-resistant, advanced, HER2-positive breast cancer (PANACEA): a single-arm, multicentre, phase 1b-2 trial. *Lancet Oncol* 2019;20:371-82.
46. Emens LA, Esteva FJ, Beresford M, et al. Trastuzumab

- emtansine plus atezolizumab versus trastuzumab emtansine plus placebo in previously treated, *HER2*-positive advanced breast cancer (*KATE2*): a phase 2, multicentre, randomised, double-blind trial. *Lancet Oncol* 2020;21:1283-95.
47. Baselga J, Bradbury I, Eidtmann H, et al. Lapatinib with trastuzumab for *HER2*-positive early breast cancer (NeoALTTO): a randomised, open-label, multicentre, phase 3 trial. *Lancet* 2012;379:633-40.
48. Bense RD, Sotiriou C, Piccart-Gebhart MJ, et al. Relevance of Tumor-Infiltrating Immune Cell Composition and Functionality for Disease Outcome in Breast Cancer. *J Natl Cancer Inst* 2016;109:djw192.

**Cite this article as:** Wen Y, Ouyang D, Chen Q, Zeng L, Luo N, He H, Anwar M, Qu L, Zou Q, Yi W. Prognostic value of tumor mutation burden and the relationship between tumor mutation burden and immune infiltration in *HER2+* breast cancer: a gene expression-based study. *Gland Surg* 2022;11(1):100-114. doi: 10.21037/gs-21-594



**Figure S1** No significant difference was observed with AJCC-TNM and pathological stages. The relationship of TMB to T (A), N (B), M (C) and stages (D). TMB, tumor mutation burden.



**Figure S2** No significant differences were observed in the infiltration levels of other immune cells between the two TMB groups. TMB, tumor mutation burden.

Table S1 DEGs between high- and low-TMB groups

Genes	logFC	P value	FDR
SMR3B	10.25878	8.00E-38	1.01E-33
INS	5.359809	7.96E-23	3.34E-19
PAGE1	5.34612	2.93E-22	9.19E-19
KRT12	5.223938	6.64E-32	4.18E-28
TBX10	3.625035	6.47E-13	1.16E-09
MAGEA4	2.95182	4.61E-06	0.000673
DCD	2.874008	1.15E-05	0.001305
ACHE	2.601691	7.20E-18	1.81E-14
TRPA1	2.340315	2.84E-07	8.93E-05
FABP7	2.30691	0.000368	0.014285
PRR4	2.302624	3.27E-07	9.56E-05
B4GALNT2	2.27809	1.39E-08	7.97E-06
FAM3D	2.116634	3.80E-07	9.95E-05
SOSTDC1	2.062557	2.07E-07	7.14E-05
ZFP57	2.044209	4.43E-06	0.000655
CLCA2	1.915096	6.37E-06	0.000837
C2CD4A	1.912346	3.54E-07	9.74E-05
MUC20	1.883634	1.16E-11	1.82E-08
CEL	1.878968	4.31E-06	0.000655
PADI3	1.789345	7.78E-06	0.000978
HOXB9	1.715189	3.03E-09	2.93E-06
ACE2	1.709743	0.000459	0.016447
NUDT8	1.699212	1.69E-11	2.36E-08
MFSD2A	1.662663	3.38E-13	7.07E-10
CYP4F22	1.643059	4.10E-06	0.000644
ABCC2	1.62889	2.54E-06	0.000432
PCP4L1	1.619974	1.33E-06	0.000279
SUSD2	1.618094	3.28E-11	4.12E-08
MUC5B	1.570947	8.56E-05	0.005541
FFAR2	1.560783	8.56E-10	8.97E-07
FIBCD1	1.551579	4.04E-07	0.000104
GSDMB	1.506932	8.89E-08	3.49E-05
CEACAM5	1.472854	0.000273	0.011875
SNPH	1.460376	1.37E-07	5.24E-05
GAL3ST2	1.457622	1.21E-05	0.001345
NUDT10	1.445567	2.10E-07	7.14E-05
HEPACAM2	1.41274	2.59E-05	0.002305
CRTAC1	1.405487	2.32E-05	0.002186
ASCL2	1.398084	3.72E-08	1.61E-05
VNN2	1.382991	5.74E-09	4.51E-06
KRT4	1.35923	0.00224	0.0455
AC131097.2	1.35631	0.000328	0.013225
NXPH4	1.332861	9.62E-06	0.001162
SLPI	1.326106	0.000425	0.015714
CALML3	1.319611	0.000968	0.027063
SOX11	1.305224	8.60E-05	0.005541
C2CD4B	1.298414	1.17E-05	0.001315
PRAC2	1.285974	0.000145	0.007868
ABCC3	1.284305	1.68E-06	0.000321
FGFR4	1.248674	1.51E-05	0.001592
ATP6V0A4	1.232072	5.11E-05	0.003711
C1QTNF2	1.220568	1.62E-08	8.86E-06
MELTF	1.197576	0.000106	0.006433
SLC13A3	1.187259	4.94E-08	2.07E-05
FUT3	1.161173	0.000309	0.012644
KRT86	1.147976	0.001942	0.041732
ITGB6	1.128633	4.11E-05	0.00329
ORMDL3	1.117312	5.20E-06	0.000727
ABCA12	1.111848	5.06E-05	0.003711
SDCBP2	1.103393	4.62E-07	0.000114
NME2	1.095368	3.86E-09	3.47E-06
AKR1C1	1.079418	0.001151	0.030834
TMEM40	1.076595	7.48E-05	0.005002
NME1-NME2	1.063632	3.36E-05	0.002777
MYCN	1.06004	0.000469	0.016528
CLEC7A	1.055267	8.95E-07	0.000194
ATP13A5	1.050073	0.000984	0.027363
MIEN1	1.039653	3.96E-05	0.003213
POF1B	1.02342	0.000285	0.012195
AIF1L	1.015252	2.75E-06	0.00046
GRB7	1.007611	0.0002	0.009455
DLX5	1.005459	0.001299	0.032426
SLC16A6	-1.00435	0.002539	0.049495
BCL2	-1.00957	0.000119	0.00704
THSD4	-1.01274	9.28E-05	0.005859
LPL	-1.02257	0.000139	0.007811
PHYHD1	-1.02559	0.000741	0.022125
IL6ST	-1.03921	2.27E-06	0.000413
CMYA5	-1.04512	3.15E-05	0.002655
IGFBP5	-1.05864	0.000247	0.01108
MS4A14	-1.06188	0.001649	0.037894
MPP2	-1.06684	0.000125	0.00725
FAM107A	-1.07561	0.002321	0.046752
REPS2	-1.07944	0.000125	0.00725
TIMP4	-1.08507	0.001292	0.032426
TRIM17	-1.09937	0.001699	0.038528
ELP2	-1.10439	1.85E-06	0.000341
KCNF1	-1.11507	0.002195	0.045179
FAM189A2	-1.1232	0.000149	0.00796
CALB2	-1.14576	0.000876	0.025069
NTRK2	-1.15761	0.000186	0.009212
MAPT	-1.17744	0.000131	0.007469
KCNE4	-1.19758	0.000698	0.021138
TUBA3D	-1.20616	0.001394	0.033691
ARTN	-1.21416	0.000236	0.010755
GREB1	-1.21593	0.000261	0.011472
AOX1	-1.24501	0.00063	0.019878
ESR1	-1.25336	0.001315	0.032497
CCDC170	-1.27815	1.24E-05	0.001364
DCDC2	-1.28487	0.000126	0.007251
SUSD3	-1.30191	0.000159	0.008397
KRT15	-1.32542	0.001482	0.035139
KCNK15	-1.3452	1.60E-05	0.001637
PARD6B	-1.3477	7.25E-07	0.000163
MMP7	-1.34811	0.00017	0.008677
AFF3	-1.36781	2.35E-05	0.002186
PRRT2	-1.39042	2.06E-05	0.002012
FABP5	-1.39257	1.42E-06	0.000287
CIDEA	-1.39277	0.001379	0.033386
SLC39A6	-1.4068	6.39E-06	0.000837
RLN2	-1.41498	0.000647	0.020078
SLC7A2	-1.42259	0.000304	0.012512
TFPI2	-1.42269	0.000329	0.013225
GFRA1	-1.44406	0.000529	0.01802
SERPINA3	-1.48629	1.47E-05	0.001561
FGF10	-1.50813	0.000231	0.010577
TFF3	-1.517	0.000343	0.013678
S100A2	-1.54293	0.00253	0.049462
EGR3	-1.55636	6.48E-07	0.000148
PXDNL	-1.57631	0.000625	0.01986
WNK4	-1.5802	0.000146	0.007868
SNORC	-1.58104	0.000725	0.021844
EN1	-1.61877	0.000102	0.006245
INSYN2	-1.64617	0.001631	0.037636
CORO6	-1.67716	0.000379	0.014514
TPRG1	-1.68165	7.23E-05	0.00488
CLIC6	-1.69964	0.000242	0.010976
PTX3	-1.70825	0.000216	0.010042
CST2	-1.72492	6.85E-06	0.000887
GABRP	-1.72599	0.001759	0.03921
KCNK2	-1.74063	0.000164	0.00861
CXCL14	-1.75219	7.21E-05	0.00488
RIMS4	-1.8241	3.61E-05	0.002966
CRISPLD1	-1.83803	2.34E-05	0.002186
CAPN6	-1.8442	0.000198	0.009435
FMO3	-1.89431	3.08E-05	0.002625
PDZK1	-1.8974	5.17E-05	0.003733
VGLL1	-1.89776	0.002086	0.043851
RGS22	-1.92686	2.34E-05	0.002186
STC2	-2.01357	3.56E-07	9.74E-05
FCRLB	-2.0321	9.38E-06	0.001156
CRABP1	-2.08692	0.000576	0.019166
IGF1R	-2.15058	1.26E-08	7.52E-06
BMPRI1B	-2.15899	4.44E-05	0.003444
TFF1	-2.16091	7.69E-05	0.005059
SCUBE2	-2.1872	8.69E-08	3.49E-05
PGR	-2.22122	2.35E-06	0.000414
SLITRK6	-2.26521	1.84E-05	0.001849
KCNJ3	-2.27728	0.000645	0.020076
CDC20B	-2.3208	0.000257	0.011451
LGALS7B	-2.39592	0.000168	0.008658
PLP1	-2.39796	0.000166	0.008627
GRPR	-2.40321	1.36E-06	0.000281
KRTDAP	-2.42576	0.000316	0.012753
ATP1A2	-2.43927	9.76E-06	0.001168
AGTR1	-2.46258	4.43E-06	0.000655
STAC2	-2.51485	4.53E-05	0.003449
SCGB3A1	-2.52054	0.000506	0.017437
HPD	-2.61451	4.04E-05	0.003257
GSTM1	-2.63201	4.50E-05	0.003449
FGL1	-2.66648	0.000645	0.020076
C20orf85	-2.71947	0.001226	0.031707
ART3	-2.72052	6.58E-05	0.004572
KRT1	-2.88843	0.000171	0.008677
PCDH10	-2.98383	0.000258	0.011472
PI15	-3.20526	2.77E-08	1.29E-05
SCGB1A1	-3.22629	0.000246	0.01108
CSN3	-3.23506	0.000507	0.017437
PZP	-3.30545	0.000111	0.00669
KRT79	-3.49986	4.67E-05	0.003539
DSG1	-3.59386	2.94E-05	0.002549
TRH	-3.93223	4.15E-05	0.003304
MUC5AC	-3.94699	5.10E-05	0.003711
RPS4Y1	-4.14079	0.000662	0.020454
CGA	-4.37307	5.59E-07	0.000135
CST5	-4.59896	7.15E-06	0.000907
S100G	-4.61529	9.54E-06	0.001162
CARTPT	-4.67963	0.000441	0.016083
FDCSP	-4.97844	2.65E-08	1.28E-05
CYP2A7	-5.16347	7.66E-07	0.000169
SERPINA6	-5.78324	6.54E-09	4.83E-06
CYP2A6	-6.43218	1.58E-10	1.80E-07

DEGs, differentially expressed genes; TMB, tumor mutation burden.

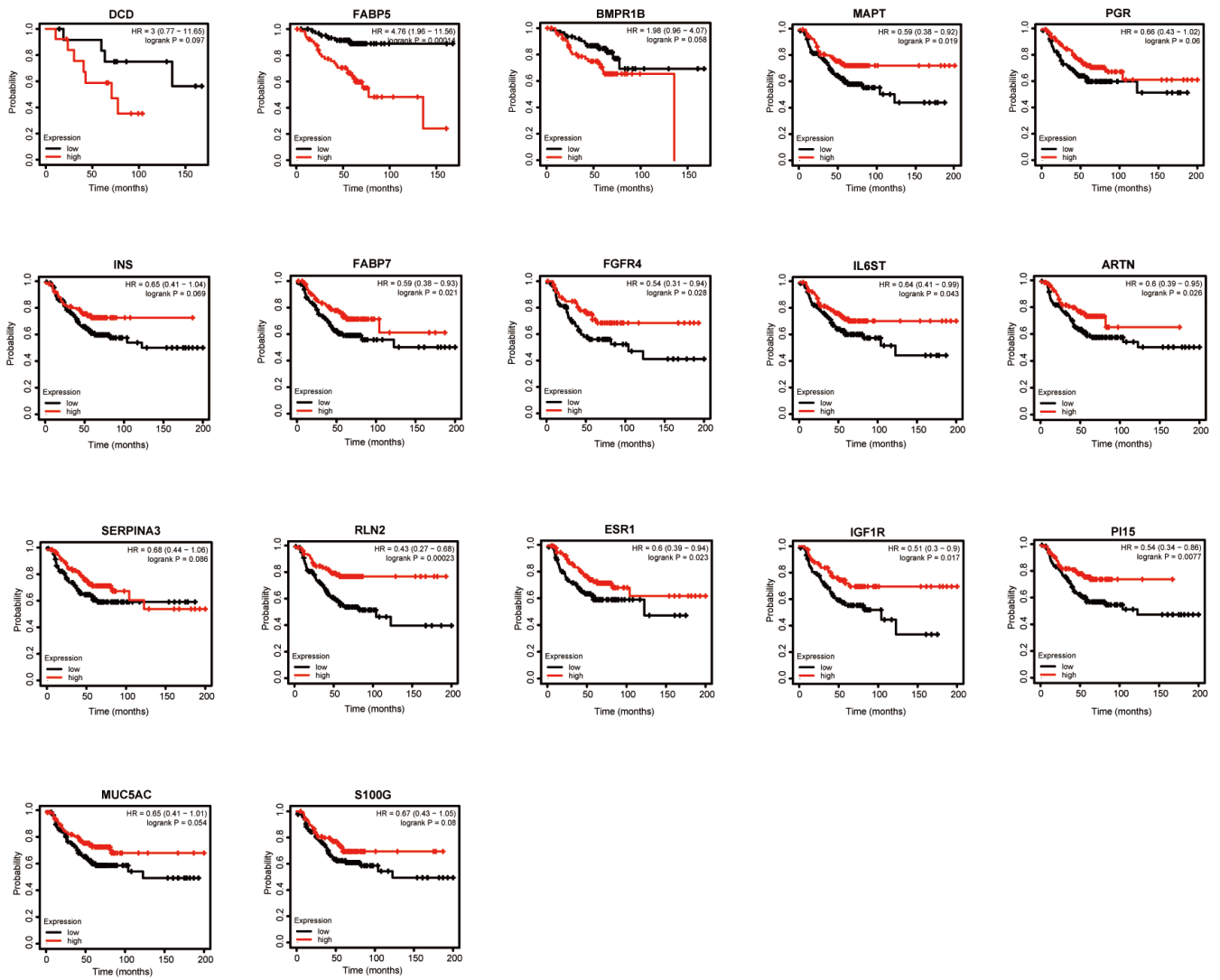


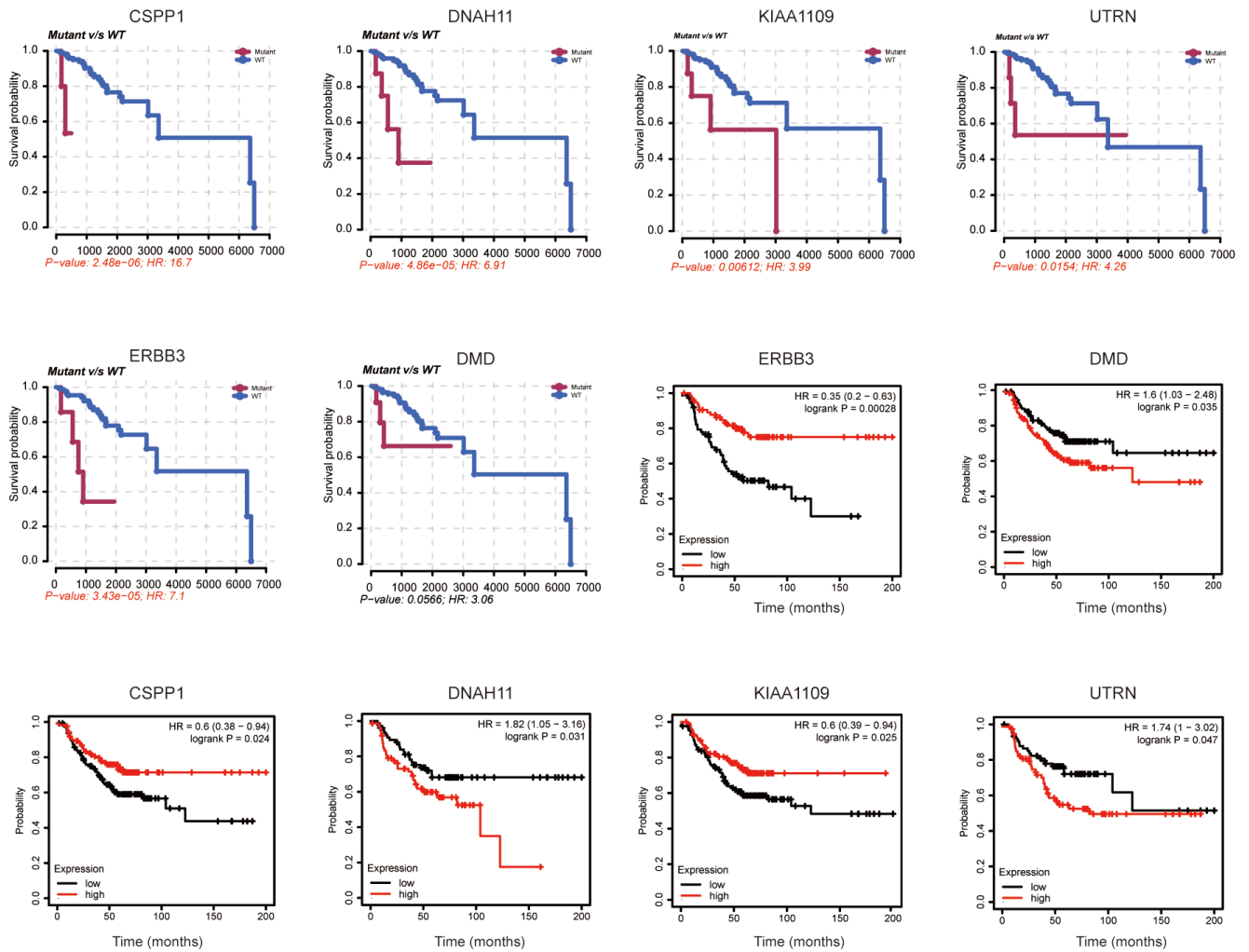
Figure S3 Survival analysis of 17 TMB-related immune genes. TMB, tumor mutation burden.



**Table S2** Differentially mutated genes between high- and low-TMB groups (top 30)

Genes	Mutation rate		P value	OR	CI. up	CI. low
	High TMB	Low TMB				
<i>TTN</i>	0.44186	0.115607	5.09E-06	5.986984	13.83298	2.620877
<i>CSPP1</i>	0.162791	0	8.17E-06	Inf	Inf	6.483162
<i>SYNE1</i>	0.209302	0.011561	9.45E-06	22.14357	218.8382	4.321598
<i>DNAH11</i>	0.186047	0.00578	1.08E-05	38.41108	1,736.446	4.894684
<i>MUC17</i>	0.255814	0.028902	1.25E-05	11.35118	44.61093	3.365185
<i>ITSN2</i>	0.139535	0	4.64E-05	Inf	Inf	5.172873
<i>PIWIL1</i>	0.139535	0	4.64E-05	Inf	Inf	5.172873
<i>ASPM</i>	0.162791	0.00578	5.55E-05	32.72031	1,502.429	4.007348
<i>CMYA5</i>	0.162791	0.00578	5.55E-05	32.72031	1,502.429	4.007348
<i>KIAA1109</i>	0.162791	0.00578	5.55E-05	32.72031	1,502.429	4.007348
<i>OTOGL</i>	0.162791	0.00578	5.55E-05	32.72031	1,502.429	4.007348
<i>HUWE1</i>	0.186047	0.017341	0.000143	12.72391	78.13288	2.876491
<i>USH2A</i>	0.186047	0.017341	0.000143	12.72391	78.13288	2.876491
<i>NCOA6</i>	0.162791	0.011561	0.000212	16.31668	167.0088	2.947262
<i>PEG3</i>	0.162791	0.011561	0.000212	16.31668	167.0088	2.947262
<i>UTRN</i>	0.162791	0.011561	0.000212	16.31668	167.0088	2.947262
<i>SYNE2</i>	0.209302	0.028902	0.000212	8.759255	35.46074	2.459223
<i>DEPDC4</i>	0.116279	0	0.000257	Inf	Inf	3.937605
<i>KAT6A</i>	0.116279	0	0.000257	Inf	Inf	3.937605
<i>KIF1B</i>	0.116279	0	0.000257	Inf	Inf	3.937605
<i>LRCH1</i>	0.116279	0	0.000257	Inf	Inf	3.937605
<i>MAST2</i>	0.116279	0	0.000257	Inf	Inf	3.937605
<i>MYLK</i>	0.116279	0	0.000257	Inf	Inf	3.937605
<i>PAPPA2</i>	0.116279	0	0.000257	Inf	Inf	3.937605
<i>RALGAPB</i>	0.116279	0	0.000257	Inf	Inf	3.937605
<i>SHANK1</i>	0.116279	0	0.000257	Inf	Inf	3.937605
<i>SZT2</i>	0.116279	0	0.000257	Inf	Inf	3.937605
<i>TG</i>	0.116279	0	0.000257	Inf	Inf	3.937605
<i>THSD7B</i>	0.116279	0	0.000257	Inf	Inf	3.937605
<i>QSER1</i>	0.139535	0.00578	0.000275	27.28709	1,280.418	3.171417

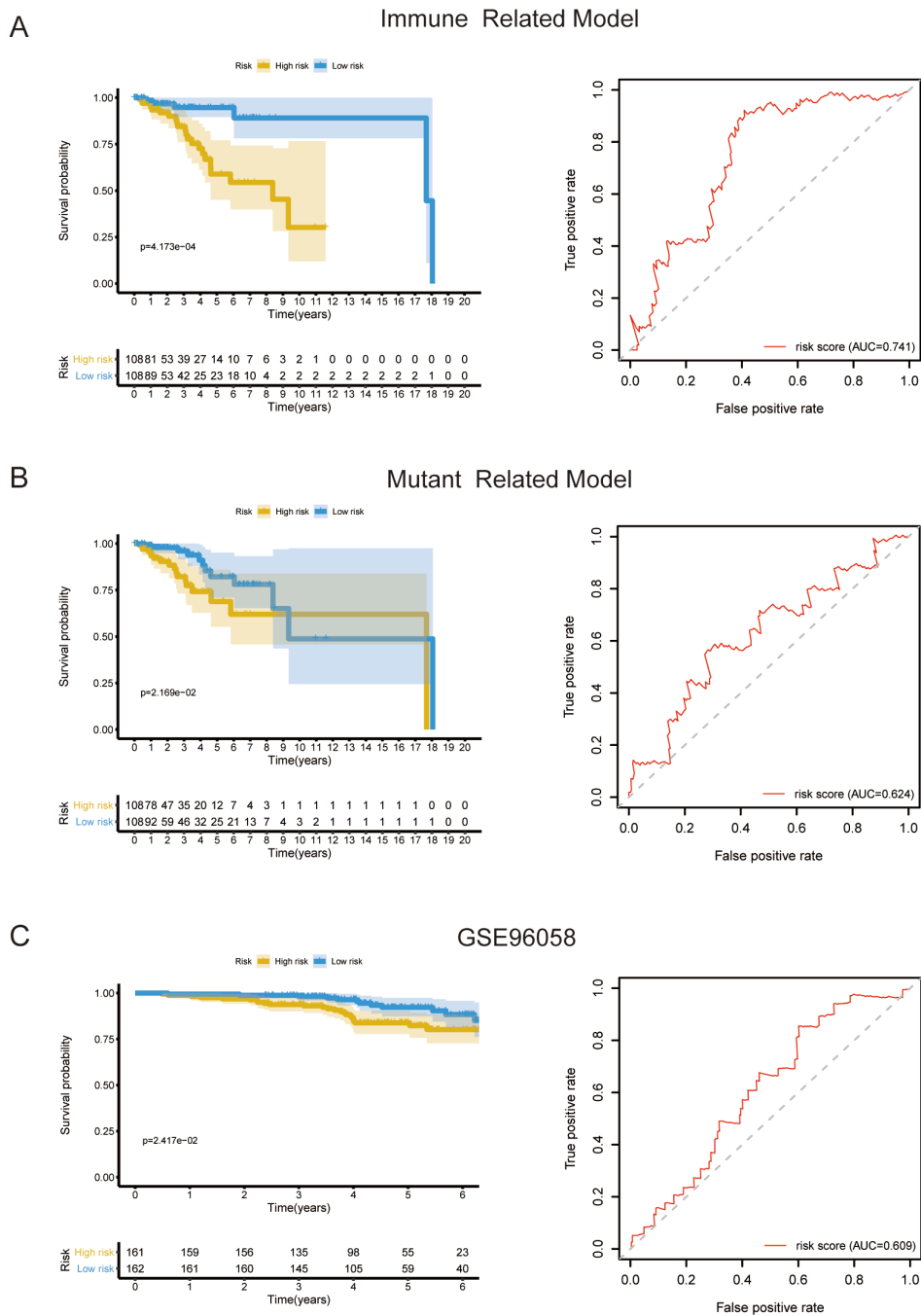
TMB, tumor mutation burden; CI, confidence interval; Inf, infinity.



**Figure S4** Mutation- and expression-related prognosis of 6 TMB-related mutant genes. TMB, tumor mutation burden.

**Table S3** The univariate Cox regression analysis of 23 genes

Genes	HR	HR.95L	HR.95H	P value
<i>ARTN</i>	0.698959	0.389483	1.254338	0.229966
<i>BMPR1B</i>	1.03753	0.869087	1.23862	0.683557
<i>CSPP1</i>	0.689729	0.365589	1.301261	0.251423
<i>DCD</i>	1.068362	0.950345	1.201035	0.268205
<i>DMD</i>	0.90872	0.417437	1.978196	0.809426
<i>DNAH11</i>	2.021264	1.029951	3.966699	0.04078
<i>ERBB3</i>	1.545548	0.933522	2.558823	0.090543
<i>ESR1</i>	0.986851	0.834388	1.167173	0.877152
<i>FABP5</i>	0.57418	0.339193	0.971963	0.038842
<i>FABP7</i>	0.982491	0.833521	1.158086	0.833241
<i>FGFR4</i>	1.191269	0.952273	1.490248	0.125543
<i>IGF1R</i>	0.846196	0.643828	1.112173	0.231082
<i>IL6ST</i>	0.916109	0.682802	1.229135	0.559046
<i>INS</i>	0.350101	0.000503	243.566	0.753296
<i>KIAA1109</i>	1.312052	0.738614	2.33069	0.354214
<i>MAPT</i>	0.693917	0.498978	0.965015	0.029886
<i>MUC5AC</i>	1.119383	0.869683	1.440778	0.381172
<i>PGR</i>	0.463558	0.229164	0.937694	0.032441
<i>PI15</i>	0.92759	0.652778	1.318095	0.674998
<i>RLN2</i>	0.878205	0.529463	1.456651	0.614929
<i>S100G</i>	1.013108	0.757935	1.354191	0.929909
<i>SERPINA3</i>	0.745016	0.413917	1.340967	0.326307
<i>UTRN</i>	0.929572	0.548193	1.576275	0.786355



**Figure S5** Identification and validation of the combined risk scoring system. (A,B) Independent immune- and mutant-related models were established, respectively. (C) Survival analysis and ROC curve of the combined model in the validation cohort (GSE96058). AUC, area under curve; ROC, receiver operating characteristic.

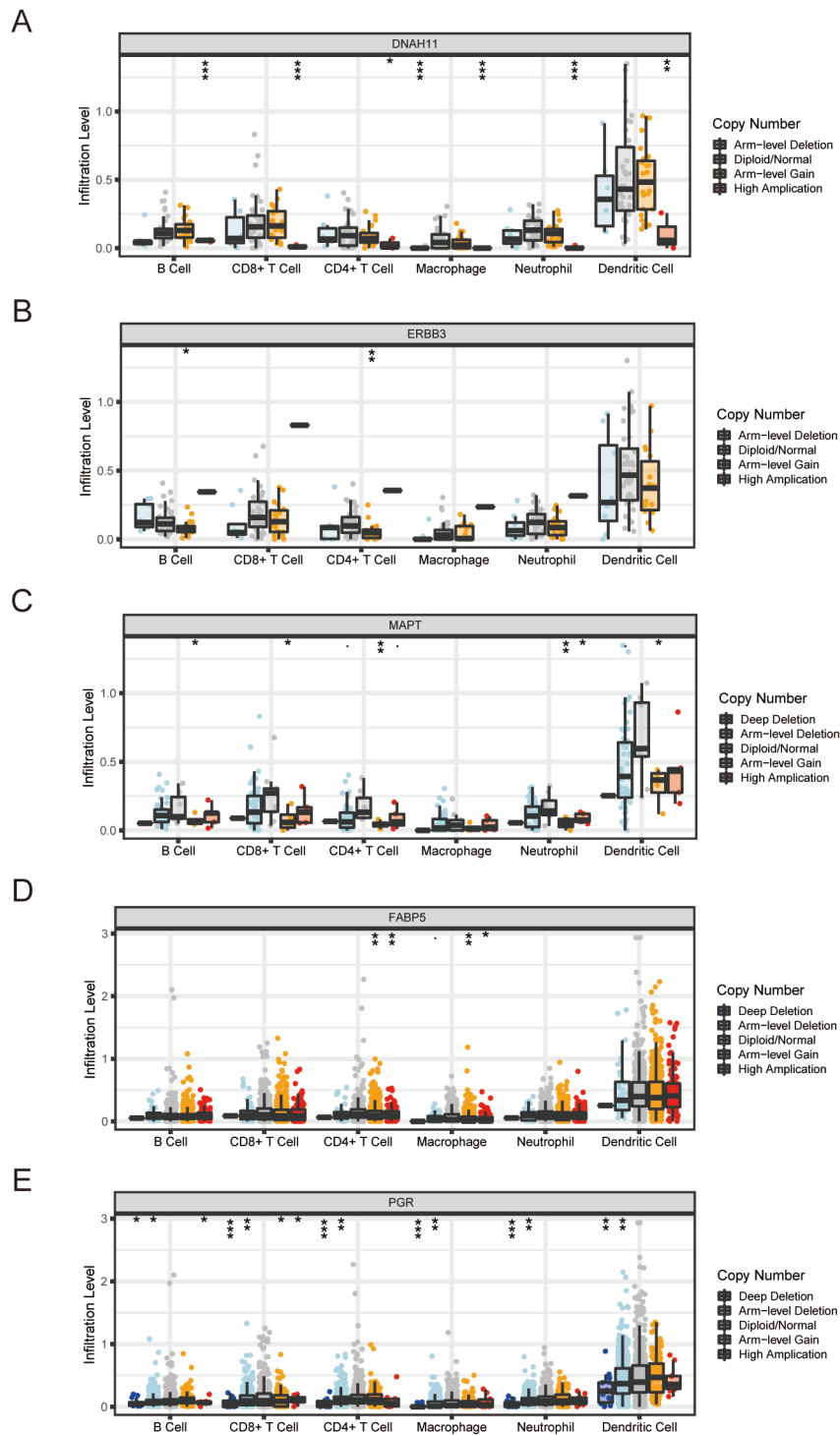
**Table S4** The immune related risk scoring system

Genes	Coef	HR	HR.95L	HR.95H	P value
<i>FABP5</i>	-0.48906	0.613205	0.365736	1.02812	0.063626
<i>FGR</i>	-0.71156	0.49088	0.238077	1.012122	0.053939
<i>MAPT</i>	-0.42131	0.656186	0.470435	0.915282	0.01309

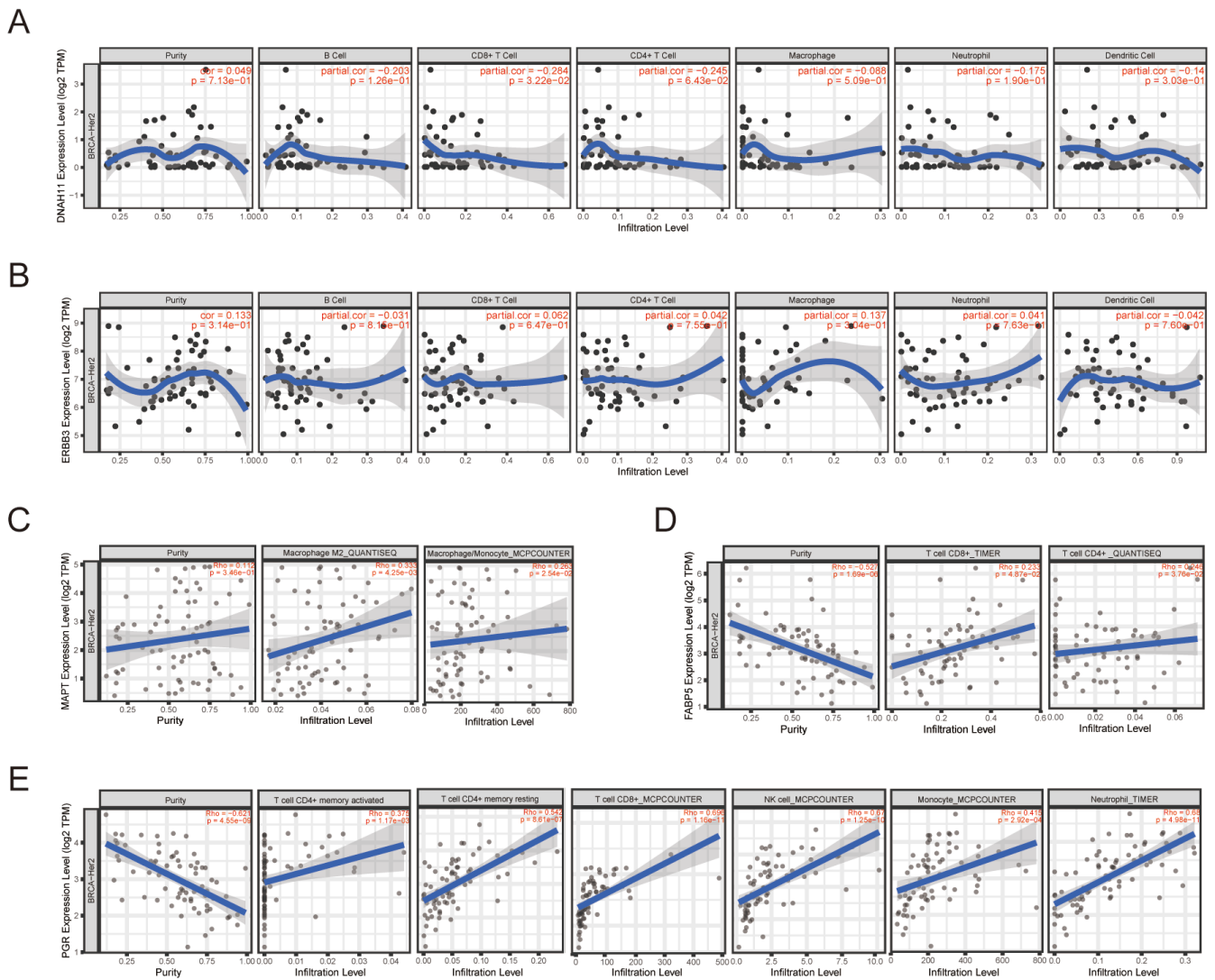
**Table S5** The mutant related risk scoring system

Genes	Coef	HR	HR.95L	HR.95H	P value
<i>CSPP1</i>	-0.21612	0.805641	0.618523	1.049365	0.109015
<i>DNAH11</i>	0.499844	1.648465	1.17012	2.322357	0.004258
<i>ERBB3</i>	0.01506	1.015174	1.001624	1.028908	0.028045





**Figure S6** Associations of 5 hub TMB-related signature mutants with immune cell infiltration. (A-E) Mutants of 5 TMB-related genes conferred the low infiltration levels of immune cells. \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ . TMB, tumor mutation burden.



**Figure S7** Correlation of the expression of 5 hub TMB-related signatures with immune infiltration levels in *HER2+* BC. (A) The expression of *DNAH11* was negatively correlated with the levels of CD8<sup>+</sup> T cells and CD4<sup>+</sup> T cells. (B) There was no significant correlation between the expression of *ERBB3* and the levels of 6 immune cells. (C) The expression of *MAPT* was associated with the levels of macrophage M2 and monocytes. (D) The expression of *FABP5* was associated with the levels of CD8<sup>+</sup> T cells and CD4<sup>+</sup> T cells. (E) The expression of *PGR* was associated with the levels of 6 immune cells. TMB, tumor mutation burden; BC, breast cancer.