

## Identification of “regulation of RhoA activity panel” as a prognostic and predictive biomarker for gastric cancer

Wenwen Huang<sup>1,\*</sup>, Songhui Zhao<sup>2,\*</sup>, Cheng Zhang<sup>1</sup>, Zhongwu Li<sup>3</sup>, Sai Ge<sup>1</sup>, Baofeng Lian<sup>2</sup>, Hui Feng<sup>4</sup>, Kai Wang<sup>2</sup>, Ruihua Xu<sup>5</sup>, Jiafu Ji<sup>6</sup>, Jing Gao<sup>7,#</sup>, Weiwei Shi<sup>2,8,#</sup>, Lin Shen<sup>1,#</sup>

<sup>1</sup>Department of Gastrointestinal Oncology, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education, Beijing), Peking University Cancer Hospital and Institute, Beijing 100142, China

<sup>2</sup>OrigiMed Inc., Shanghai 201112, China

<sup>3</sup>Department of Pathology, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education, Beijing), Peking University Cancer Hospital and Institute, Beijing 100142, China

<sup>4</sup>Shanghai Junshi Biosciences Co., Ltd, Shanghai 201203, China

<sup>5</sup>State Key Laboratory of Oncology in South China, Department of Medical Oncology, Sun Yat-sen University Cancer Center, Collaborative Innovation Center for Cancer Medicine, Guangzhou 510060, China

<sup>6</sup>Department of Gastrointestinal Surgery, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education, Beijing), Peking University Cancer Hospital and Institute, Beijing 100142, China

<sup>7</sup>National Cancer Center, National Clinical Research Center for Cancer, Cancer Hospital and Shenzhen Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen 518116, China

<sup>8</sup>Department of Thoracic Surgery, Shanghai Pulmonary Hospital, School of Medicine, Tongji University, Shanghai 200433, China

\*Equal contribution

#Senior co-authors

**Correspondence to:** Lin Shen, Weiwei Shi, Jing Gao; **email:** [shenlin@bjmu.edu.cn](mailto:shenlin@bjmu.edu.cn), [shiw@origimed.com](mailto:shiw@origimed.com), [gaojing\\_pumc@163.com](mailto:gaojing_pumc@163.com), <https://orcid.org/0000-0002-3613-9413>

**Keywords:** regulation of RhoA activity, gastric cancer, prognosis and predictive biomarker, tumor migration, tumor microenvironment

**Received:** February 25, 2020

**Accepted:** September 14, 2020

**Published:** December 03, 2020

**Copyright:** © 2020 Huang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/3.0/) (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### ABSTRACT

RhoA is a member of the RHO family GTPases and is associated with essential functions in gastric cancer. In this study, we identified a gastric cancer biomarker, termed the “regulation of RhoA activity panel” (RRAP). Patients with gastric cancer from The Cancer Genome Atlas database were divided into training (N=160) and validation (N=155) cohorts. A cohort of 109 Chinese gastric cancer patients was utilized as an independent validation. Patients with mutated RRAP showed significantly better overall survival than patients with wild type RRAP. We also analyzed the association between RRAP and the migration capacity, immune-related signatures, and the tumor microenvironment. RRAP-mutant tumors had a significantly lower degree of lymph node metastasis and lower activities of migration-related pathways. These tumors also showed significantly increased immune cell infiltration and cytotoxic activity. Furthermore, two independent patient cohorts who received immune checkpoint blockade therapy were assessed for RRAP mutant status. As expected, for both immunotherapy cohorts, higher response rates to immune checkpoint blockade therapy were observed in patients with RRAP-mutant tumors than in patients with wild type RRAP tumors. Overall, this study indicates that the RRAP gene set is a potential biomarker for gastric cancer prognosis and therapeutic selection.

## INTRODUCTION

Gastric cancer is one of the most frequently occurring and lethal malignancies [1]. The Lauren classification divides gastric cancer into 2 main subtypes, intestinal and diffuse [2]. However, this classification provides limited guidance for disease prognosis and treatment decisions. Global efforts to characterize gastric cancer at the molecular level from the perspective of cancer genomics and transcriptomics have been made, including those from The Cancer Genome Atlas (TCGA) [3] and the Asian Cancer Research Group [4]. Gastric cancer patients are classified into 4 subtypes based on gene expression profiling; each subtype exhibits distinct patterns of molecular alterations, disease progression, and prognosis [4]. Although these large-scale efforts have provided comprehensive insights into gastric cancer, they have not translated into a clinical benefit. A genomic-based molecular biomarker with prognostic and/or therapy predictive value is still needed for gastric cancer.

Among the various genomic alternations that occur in gastric cancer, *RHOA* mutation plays a critical role in the development and progression of cancer by regulating actin organization [5], cell migration [6], cytokinesis and the cell cycle [7]. Recent studies have also suggested its potential role in modulating the tumor microenvironment (TME) of cancers [8, 9]. Although the overexpression of RhoA has been frequently recognized in various cancers and was found to be significantly associated with poor prognosis in gastric cancer [10], similar overall survival (OS) rates were nevertheless observed between patients with *RHOA* mutant and wild-type gastric cancers [11, 12]; therefore, the prognostic value of this gene mutation is poor. In 2017, Shi et al. [13] established that mutations within the gene set designated as the “regulation of RhoA activity pathway” were associated with better progression-free survival (PFS) and overall survival in HER2+ breast cancer patients. This gene set includes RhoA, as well as guanine nucleotide exchange factors and GTPase-activating proteins [14], both of which are involved in regulating RhoA activity. Furthermore, altered RhoA signaling has been reported in gastric cancers, especially in the diffuse type gastric cancer [3, 15–21]. However, the clinical significance of this pathway in overall gastric cancer remains unresolved. Mutations in “regulation of RhoA activity pathway” gene set may be involved in the prognosis and therapeutic prediction of gastric cancer through their effect on RhoA and its effector molecule activity.

Based on the pathway changes of RhoA activity, we developed a statistically optimized gene subset as a biomarker by applying a genetic algorithm to a training

gastric cancer cohort obtained from the gastric cancer dataset of TCGA. We validated this biomarker in a nonoverlapped TCGA-validation cohort and in an independent Chinese gastric cancer (CGC) cohort. The association of this biomarker with lymph node metastasis, migration-related pathways, immune-related signatures, and the TME was assessed to glean insights into possible related mechanisms. Inspired by its effect on the TME, we collected 2 independent cohorts of gastric cancer patients (denoted as IM1 and IM2) who received immune checkpoint blockade (ICB) treatment and revealed the potential predictive capability of this biomarker.

## RESULTS

### Prognostic biomarker “regulation of RhoA activity panel” for gastric cancer

As shown in Figure 1, the potential biomarker “regulation of RhoA activity panel” (RRAP) was calculated from the TCGA training cohort with a genetic algorithm on the “regulation of RhoA activity pathway” gene set (see Methods). The resulting optimal solution, containing 20 genes, is hereafter denoted as RRAP (Figure 1B). These gene mutations were significantly associated with gastric cancer (Supplementary Table 1). According to whether or not mutations occurred in the coding region of any of the RRAP genes, patients were classified as RRAP-wild type or RRAP-mutant. In the TCGA training cohort, RRAP-mutant patients displayed better OS compared to RRAP-wild type patients, with a hazard ratio of 0.4 (95% confidence interval: 0.2-0.79, p-value = 0.006, Figure 2A). Importantly, this was also validated in the TCGA validation cohort, with a hazard ratio of 0.48 (95% confidence interval: 0.26-0.91, p-value = 0.021, Figure 2B). Furthermore, we incorporated age, gender, Lauren classification, pathologic stage, and RRAP into a multivariate Cox analysis for overall survival. The results showed that RRAP was an independent prognostic factor after adjusting for these clinicopathological parameters in both the TCGA training (hazard ratio = 0.35, 95% confidence interval: 0.17-0.7, p-value = 0.003; Supplementary Figure 1A) and TCGA validation cohorts (hazard ratio = 0.51, 95% confidence interval: 0.26-1, p-value = 0.05; Supplementary Figure 1B).

The reliability of RRAP was assessed using an independent patient cohort that was a subset of the 109 CGC patients. These patients were whole-exome sequenced, and 26 (23.9%) were identified as RRAP-mutant. This cohort exhibited a mutational proportion similar to the TCGA training (26.3%) and TCGA validation (28.3%) cohorts. For the reliability

assessment cohort, Kaplan-Meier curves showed that RRAP-mutant was significantly associated with longer OS (hazard ratio = 0.13, 95% confidence interval: 0.03-0.52, p-value = 6.66e-4, Figure 2C). Multivariate Cox regression analysis was performed as described above, and the results again showed that RRAP was an independent prognostic factor (Supplementary Figure 1C). We also merged data of the TCGA training, TCGA validation, and CGC cohorts into overall cohort. RRAP-mutant patients displayed better OS compared to RRAP-wild type patients, with a hazard ratio of 0.38 (95% confidence interval: 0.25-0.59, p-value = 7.37e-6) in the overall cohort (Figure 2D and Supplementary Figure 1D).

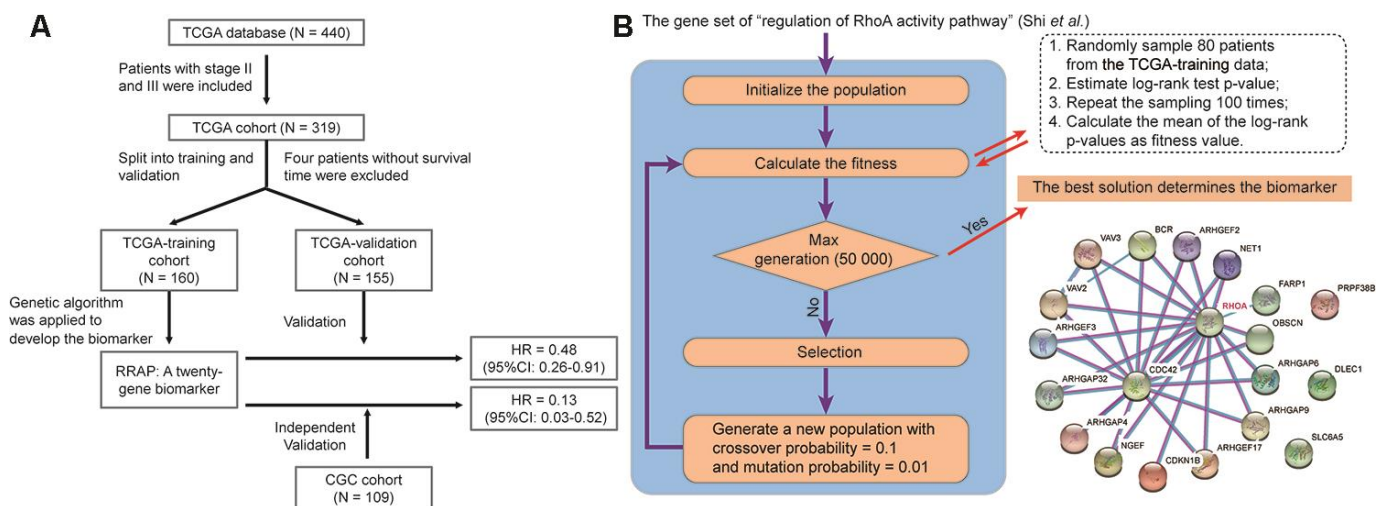
### RRAP was associated with cell migration activity

Given the role of RhoA in cell migration, we assessed the effect of RRAP on tumor metastasis. Strikingly, patients with RRAP-mutant tumors had a significantly lower rate of lymph node metastasis (pN3 of regional lymph node) compared with RRAP-wild type patients (11.5% vs 36.1%, respectively, Fisher's exact test p-value = 0.026) in the CGC cohort, which was also confirmed in the overall TCGA cohort (12.8% vs 24.7%, respectively, p-value = 0.03). We also found that patients with RRAP-mutant tumors had a moderately lower distant metastasis recurrence risk compared to patients with RRAP-wild type tumors in the CGC cohort (15.4% vs 31.3%, respectively, p-value = 0.136); the same trend was observed in the overall TCGA cohort (3.6% vs. 7.9%, respectively, p-value = 0.26). This suggested that RRAP-mutant tumors may have an impaired migration capacity. At the molecular level, we explored the RNA-seq data collected from

TCGA and estimated the enrichment score of 4 migration-related functions—adherens junctions, cell adhesion molecules, focal adhesion, and regulation of the actin cytoskeleton. All of these functions showed significantly lower activity in RRAP-mutant tumors compared to RRAP-wild type tumors (Figure 3A–3D). Further analysis revealed that 39 genes related to the 4 functions exhibited significantly lower expression in RRAP-mutant tumors (log2 fold change > 1 and adjusted p-value < 0.01, Figure 3E). We next performed univariate Cox and Kaplan-Meier analysis on the 39 genes and found that 5 genes (*CLDN11*, *CLDN6*, *CLDN9*, *VTN*, and *F2*; Supplementary Table 2) were significantly associated with poor OS (Figure 3F, log-rank p-value < 0.01). Kaplan-Meier analysis revealed that the low expression of these 5 genes, along with *CNTN1*, *CNTN2*, *CADM1*, *NCAM1*, *FGF19*, and *FGF20*, were significantly associated with better OS (Figure 3G, Supplementary Figure 2). Taken together, these results indicated an association between RRAP mutation and tumor cell migration, thus resulting in metastasis.

### Association of RRAP with immune-related signatures

We next explored the correlation between RRAP and the immune-related signatures of tumor mutational burden (TMB), neoantigen burden (NAB), and deficient mismatch repair (dMMR). Higher TMB was observed in RRAP-mutant patients in the TCGA training, TCGA validation, and CGC cohorts, with p-values of 3.7e-8, 9e-11, and 6.3e-5, respectively (Figure 4A). A significantly higher NAB (available only in the CGC cohort) was observed in RRAP-mutant patients (p-value

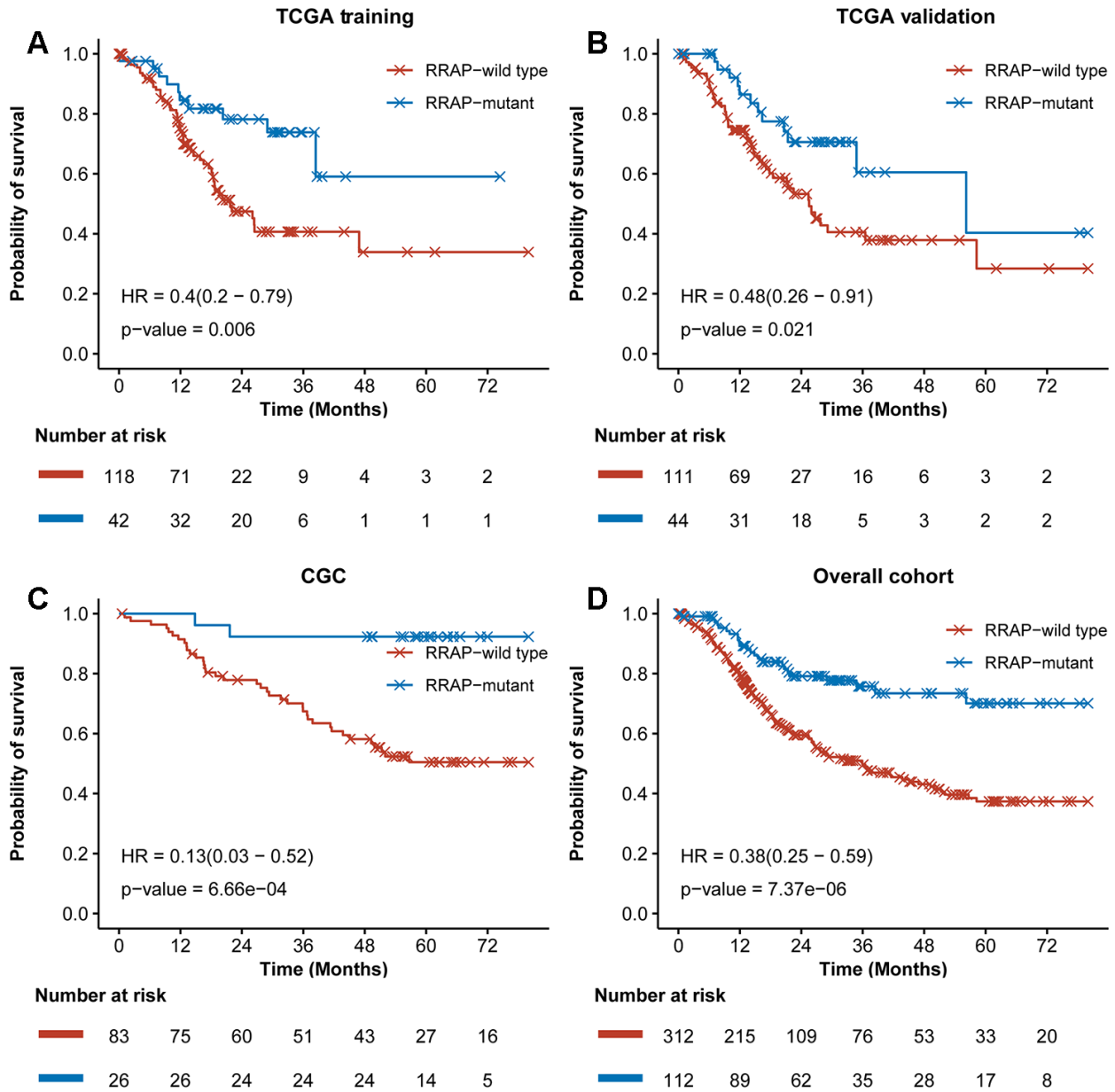


**Figure 1. Identification of RRAP biomarker.** (A) Outline of the cohort construction and analysis workflow. (B) RRAP selection. The right panel shows the interaction of genes among RRAP.

= 0.002, Figure 4B). Additionally, there was a significantly higher fraction of mutational signatures associated with dMMR in the RRAP-mutant tumors compared to RRAP-wild type tumors in all 3 cohorts (Figure 4C). Although the 3 immune-related signatures in this study could predict the response to immunotherapy in multiple cancers [22], TMB and NAB were not associated with the overall survival of gastric cancer patients, while dMMR was related to overall survival in the TCGA validation cohort (Figure 4D).

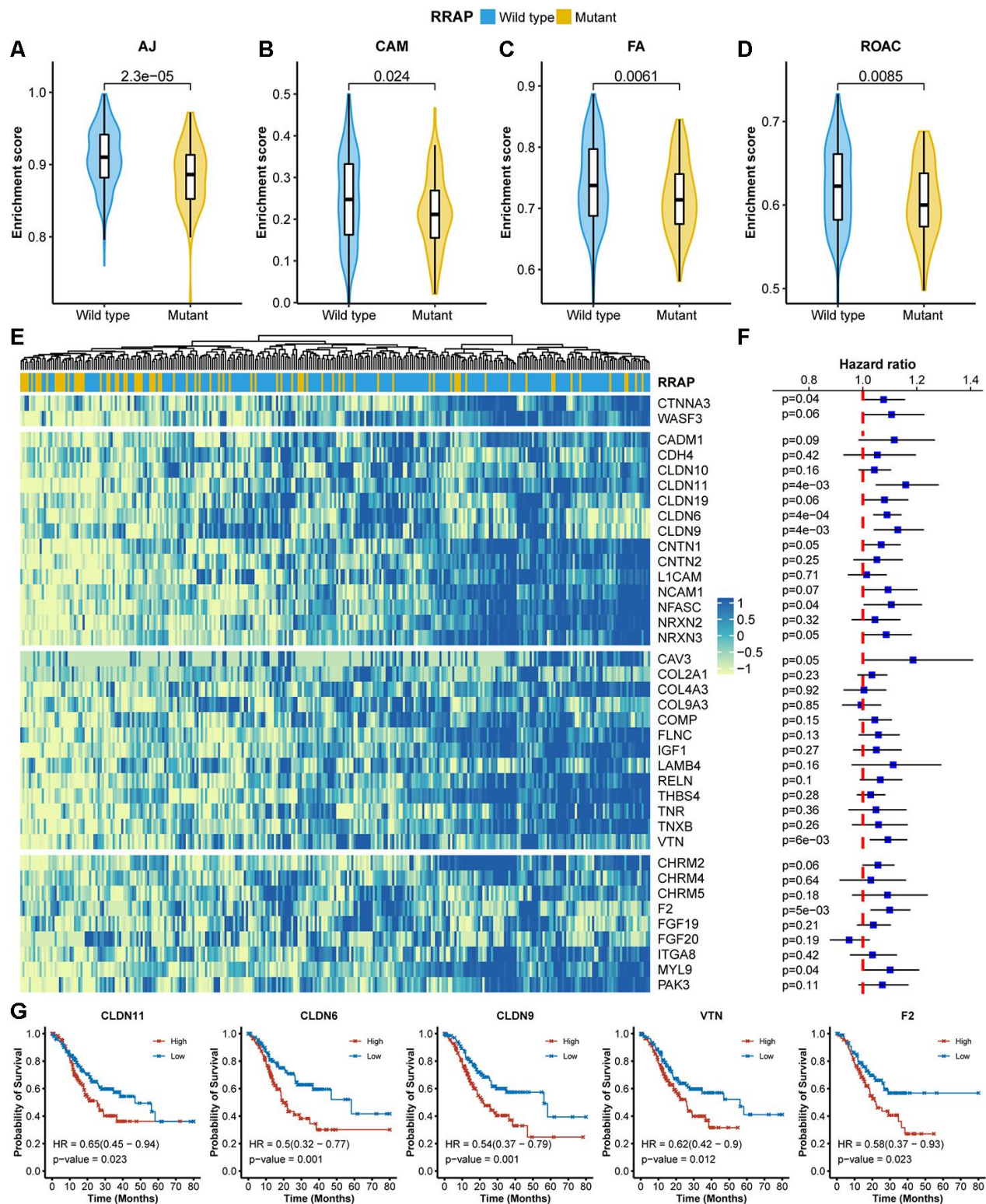
### Effect of RRAP mutation on the tumor microenvironment

In addition to the 3 signatures above, we further explored the association between RRAP and the TME using TCGA expression data. For this purpose, we collected tumor immune infiltrate data published by TCGA and compared immune cellular fraction between RRAP-mutant and RRAP-wild type tumors [23]. RRAP-mutant tumors showed significantly higher levels of infiltration of activated CD4<sup>+</sup> memory T cells



**Figure 2. Training and validation of RRAP.** Kaplan-Meier survival curves for RRAP-wild type and RRAP-mutant patients in the (A) TCGA training cohort, (B) TCGA validation cohort, (C) an independent CGC cohort, (D) and the overall cohort containing merged data from the three cohorts. P-values were estimated with the log-rank test, and hazard ratios (HRs) were estimated with the Cox model.



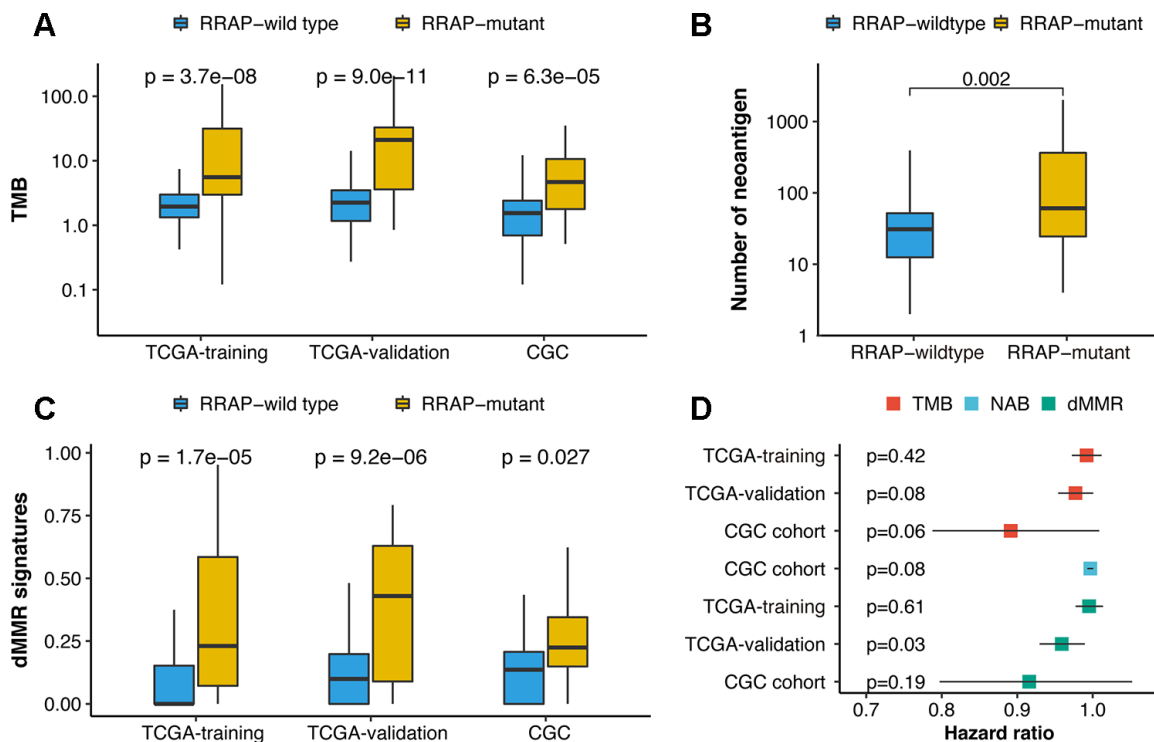


**Figure 3. Association of RRAP with lower cell migration activity.** (A–D) Pathway enrichment scores between RRAP-wild type and RRAP-mutant were compared among 4 functions: adherens junction (AJ, A), cell adhesion molecules (CAM, B), focal adhesion (FA, C), and regulation of the actin cytoskeleton (ROAC, D). The Wilcoxon rank-sum test was applied to estimate differences. (E) Heatmap of differentially expressed genes among the 4 functions; the column was clustered and annotated with RRAP status. (F) Forest plot of association of between gene expression and overall survival. Squares indicate the hazard ratios and error bars represent the 95% confidence interval; the log-rank test was performed to estimate p-values. (G) Kaplan-Meier overall survival curves for patients with high and low expression levels of CLDN11, CLDN6, CLDN9, VTN, and F2. P-values were estimated with the log-rank test, and HRs were estimated with the Cox model.

(p-value = 8.6e-7) and M1 macrophages (p-value = 0.00077) (Figure 5A), both of which are reportedly associated with longer survival in multiple cancers [24, 25]. In addition, we observed higher infiltration of CD8+ T cells—the primary effector in anti-tumor immunity—in RRAP-mutant tumors compared to RRAP-wild type tumors, although it did not reach a statistical significance (p-value = 0.07385, Figure 5A). There was a significantly difference in the levels of CD8+ infiltration in the independent CGC cohort. IHC staining was used to evaluate the infiltration of CD8+ tumor infiltrating lymphocytes (TILs) in 52 tumors from the CGC cohort; 40 of these were RRAP-wild type tumors and 12 were RRAP-mutant tumors. Consistent with the TCGA RNAseq results, RRAP-mutant tumors exhibited increased infiltration of CD8+ TILs compared to RRAP-wild type tumors (p-value = 0.026) (Figure 5B–5F). Given the functional dependence of CD4+ T cells and M1 macrophages on human leukocyte antigen (HLA) class II molecules, further analyses of several HLA class II molecules were also performed. The results showed significantly higher expression of these molecules in RRAP-mutant tumors compared with RRAP-wild type tumors (Figure 5G), including HLA-

DMA (p-value = 0.0106), HLA-DQA1 (p-value = 0.0235), and HLA-DRA (p-value = 0.0091). Taken together, an increased anti-tumor immune response by CD4+ T cells and macrophages could be seen for RRAP-mutant tumors.

Granule exocytosis-related molecules (e.g., PRF1/GZMA/GZMB) and death ligand pathways (e.g., the Fas/FasL apoptotic killing pathway) are involved in the cytotoxic effects of CD8+ T cells [26] and possibly of CD4+ T cells [27, 28]. We found that PRF1, GZMA, and GZMB were significantly increased in RRAP-mutant tumors compared to RRAP-wild type tumors, with p-values of 0.0056, 0.002, and 4.5e-5, respectively (Figure 5H). Additionally, RRAP was associated with significantly increased FAS (p-value = 0.0039) and FASLG expression (p-value = 0.0207) (Figure 5I). The apoptosis pathway was also evaluated and exhibited higher activity in RRAP-mutant tumors than in RRAP-wild type tumors (p = 0.0014, Figure 5J). These results strongly suggested that RRAP mutation affected the TME, providing evidence for the association of RRAP-mutant with better OS on the basis of increasing anti-tumor activity in the TME (Supplementary Figure 3).

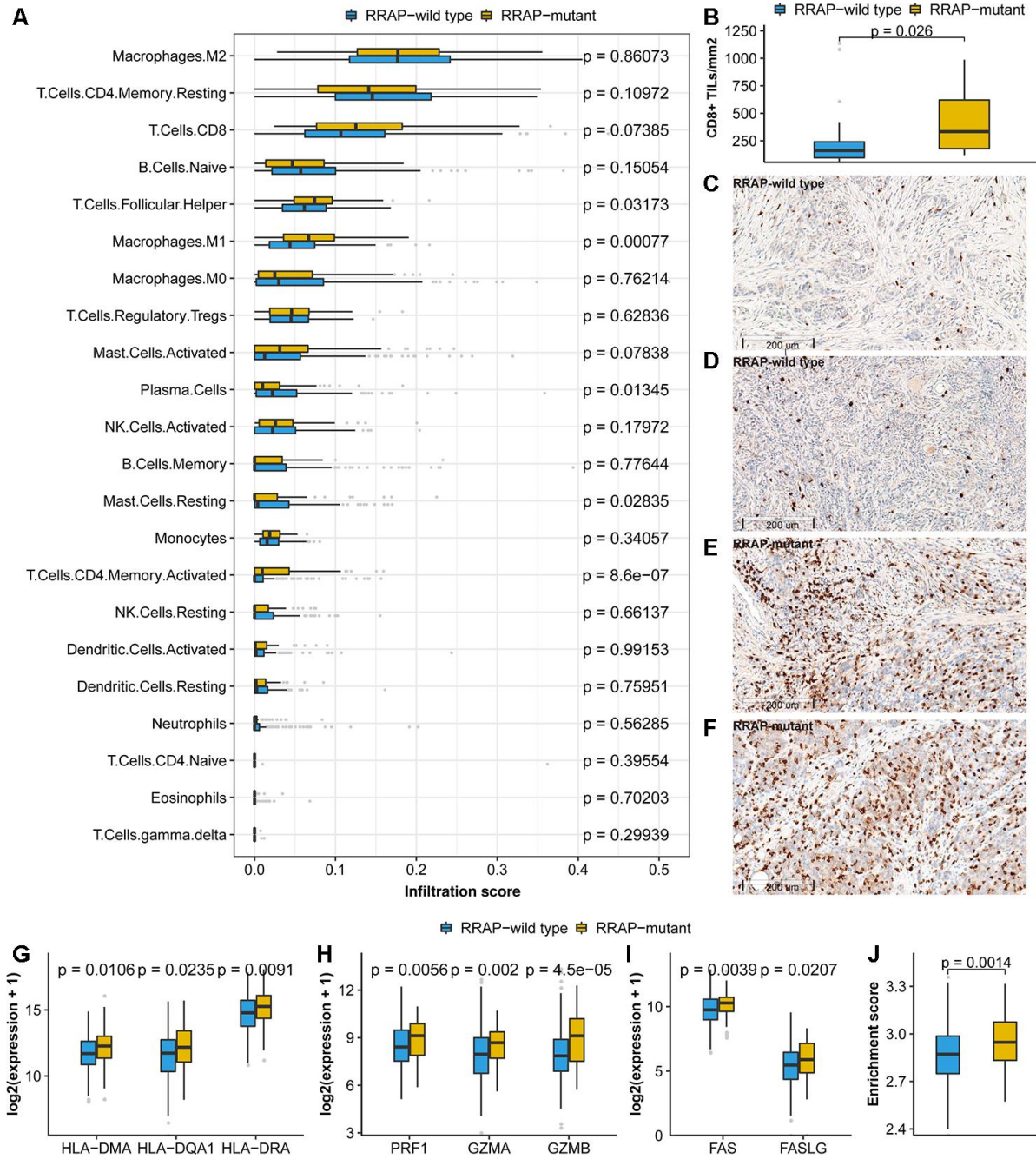


**Figure 4. Association of RRAP with immune-related biomarkers.** (A) Boxplot for differences in the tumor mutational burden (TMB) between RRAP-wild type and RRAP-mutant tumors in the TCGA training, TCGA validation, and CGC cohorts. (B) Boxplot for differences in the neoantigens between RRAP-wild type and RRAP-mutant tumors in the CGC cohort. (C) Boxplot for differences in the dMMR percentage between RRAP-wild type and RRAP-mutant tumors in TCGA training, TCGA validation, and CGC cohort. The Wilcoxon rank-sum test was applied to compare the differences. (D) Association of overall survival and TMB, neoantigen burden (NAB), and dMMR in the 3 cohorts. The hazard ratio was estimated with univariate Cox analysis, and the log-rank test was applied to calculate the p-value.

## RRAP as a potential predictive biomarker for checkpoint inhibitor-based immunotherapy

Inspired by the significant effect of RRAP on prognosis, immune-related signatures and the TME, we further

investigated the association between RRAP and the response to ICB therapy. Of the IM1 cohort who received ICB therapy (N = 37), 7 patients were identified as RRAP-mutant and 30 were RRAP-wild type. After treatment, 15 (50%) patients in the RRAP-



**Figure 5. Association of RRAP with the tumor microenvironment.** (A) Immune infiltrations estimated by CIBERSORT were compared between RRAP-wild type tumors and RRAP-mutant tumors. (B) The density of CD8+ TILs was compared between RRAP-wild type and RRAP-mutant tumors. (C–F) The representative immunohistochemistry for CD8 images of RRAP-wild type tumor tissue (C, D) and RRAP-mutant tumor tissue (E, F). (G–I) Expression values were compared for HLA class II genes (G), cytotoxic effector molecules (H), and apoptosis-related genes (I). (J) The boxplot shows the difference in enrichment score of the apoptosis pathway between RRAP-wild type and RRAP-mutant tumors. The Wilcoxon rank-sum test was applied to estimate the p-values.



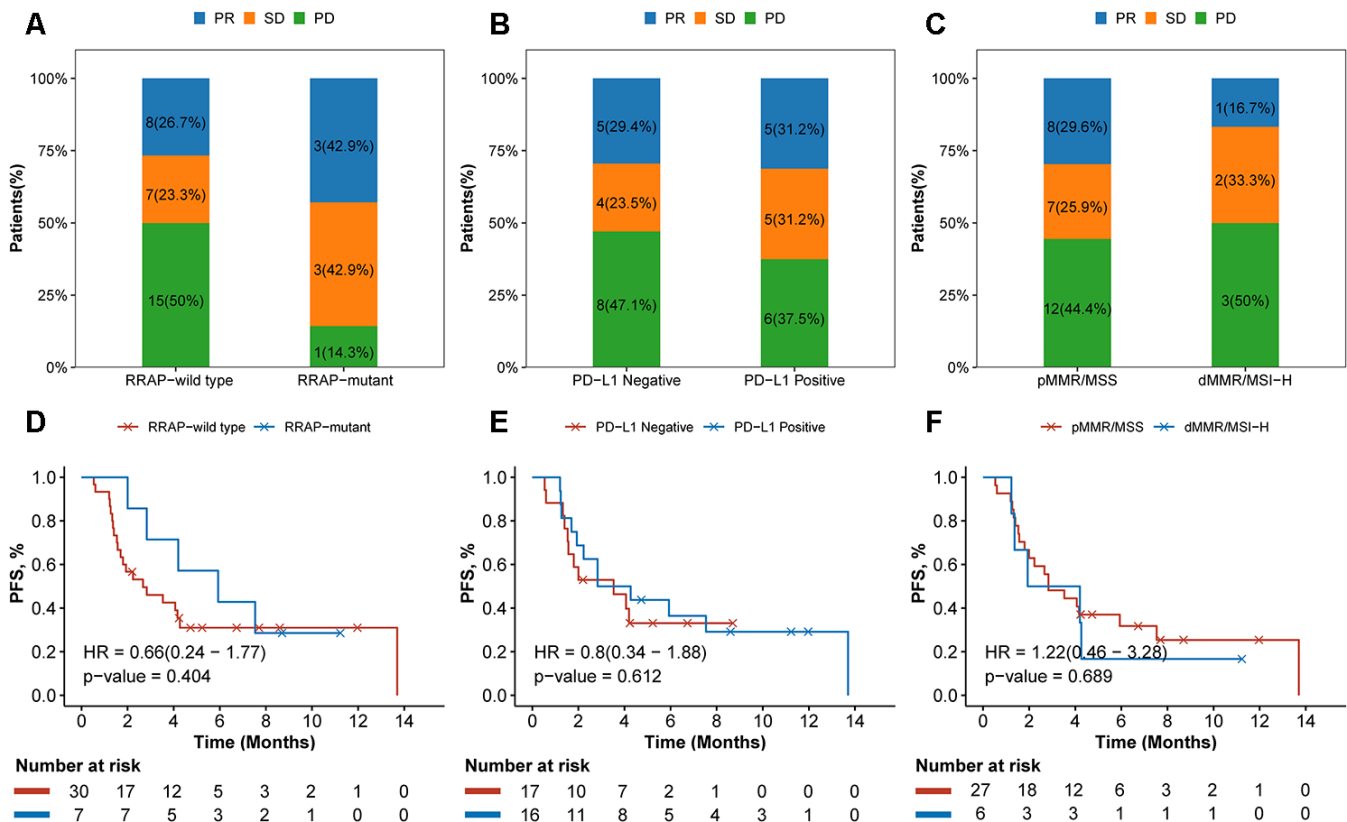
wild type group had progressive disease. In contrast, only 1 (14.3%) patient in the RRAP-mutant group had progressive disease (Figure 6A). We also evaluated the efficacy of ICB therapy between groups stratified by PD-L1 expression and microsatellite instability (MSI) and/or mismatch repair (MMR) status, and no notable differences were found (Figure 6B, 6C). The Kaplan-Meier curves suggested that RRAP-mutant favored PFS compared with RRAP-wild type (Figure 6D), with respective median PFS times of 5.93 months (95% confidence interval: 2.83-not available) and 2.67 months (95% confidence interval: 1.70-not available). However, there was no difference in PFS between PD-L1-positive and PD-L1-negative patients (median PFS: 3.55, 95% confidence interval: 1.93-not available and 3.53, 95% confidence interval: 1.57-not available, respectively) (Figure 6E). The same observation was made when comparing dMMR/MSI-H and pMMR/MSS patients (median PFS: 3.07 95% confidence interval: 1.37-not available and 2.83, 95% confidence interval: 2-not available, respectively) (Figure 6F). In the IM2 cohort (N = 47), 12 patients were classified as RRAP-mutant; these patients had a better overall response rate

than RRAP-wild type patients (33.3% vs 8.6%, respectively, Fisher's exact test p-value = 0.06). These results suggested that RRAP may play a role as a potential predictive biomarker for checkpoint inhibitor-based immunotherapy in gastric cancer.

## DISCUSSION

In this study, we identified RRAP as a biomarker, validated its prognostic effect in TCGA and CGC gastric cancer data sets, and investigated its association with tumor metastasis, the TME, and its potential prediction value for ICB therapies.

For the TCGA training, TCGA validation, and CGC cohorts, the frequencies of RRAP-mutant were elevated (26.3%, 28.3%, and 23.9%, respectively) compared to *RHOA* mutations alone (7.5%, 4.5%, and 3.7%, respectively), suggesting that RRAP can be more broadly utilized as a predictive biomarker for gastric cancer. RRAP-mutant proved to be an independent prognostic factor and was significantly associated with favorable overall survival, regardless of histological



**Figure 6. Association of RRAP and immunotherapy efficacy.** (A–C) Stacked barplot showing the proportion of patients with progressive disease (PD), stable disease (SD), or partial response (PR) for each group divided by RRAP (A), PD-L1 (B) and MSI/MMR status (C). (D–F) The text indicates the number and percentage of patients in each group. Kaplan-Meier survival curves for patients grouped by RRAP (D), PD-L1 (E), and MSI/MMR status (F). P-values were estimated with the log-rank test, and HRs were estimated with the Cox model.



classification (i.e., not only in diffuse gastric cancer). This clinical significance is likely attributed to altered RhoA activity disrupting the process of tumor invasion and metastasis when the process is impaired by RRAP mutation. Clinically, our results showed that patients with RRAP-mutant tumors had a significantly lower rate of lymph node metastasis (pN3) and a lower risk of distant metastasis recurrence. At the molecular level, our analysis suggested that RRAP-mutant tumors exhibited a low activity of migration-related functions (adhesion junctions, cell adhesion molecules, focal adhesion, and regulation of the actin cytoskeleton). In these functions, 11 genes (*CLDN11*, *CLDN6*, *CLDN9*, *VTN*, *F2*, *CNTN1*, *CNTN2*, *CADM1*, *NCAM1*, *FGF19*, and *FGF20*) that were downregulated in RRAP-mutant tumors were significantly associated with improved OS. Our findings are supported by previous results, not only regarding the role of RhoA in tumor cell invasion and metastasis [29] but also regarding the association of RhoA activity with gastric cancer prognosis [30, 31].

Recent efforts have shown that RhoA can modulate the TME [8, 9]. Therefore, we further assessed the significant association between RRAP and clinical outcomes in terms of the TME. Our results indicated that RRAP-mutant patients had higher fractions of activated CD4+ memory T cells, CD8+ T cells, and M1 macrophages, all of which have been reported as biomarkers that are positively associated with overall survival in multiple cancers [24, 25]. We further found that HLA-DMA, HLA-DQA1, and HLA-DRA, on which CD4+ T cells and M1 macrophages functionally depend [32, 33], also showed significantly increased expression in RRAP-mutant tumors. Our experimental CD8+ IHC results were highly consistent with our bioinformatics analysis. Moreover, we also showed that mutated RRAP may affect the TME by regulating the expression levels of granule exocytosis-related molecules (PRF1/GZMA/GZMB) and the death ligand pathway (Fas/FasL apoptotic killing pathway), which are involved in the cytotoxic effects of CD8+ T cells and possibly of CD4+ T cells. Taken together, these results strongly indicated an association between RRAP-mutant and better OS on the basis of increasing antitumor activity in the TME (Supplementary Figure 3). The association between well-known immunotherapy-related biomarkers (TMB, NAB, and dMMR) and the RRAP status was also assessed. Interestingly, all three biomarkers were significantly higher in RRAP-mutant tumors; however, only RRAP had prognostic significance. Moreover, RhoA signaling plays an important role in inducing activating innate immune and adaptive T cell responses [34]. For example, the downregulation of CDC42 reduces NK cell-mediated killing, allowing cancer cells to escape from the human immune response [35]. VAV3, a Rho

family GTPase, activates multiple cell signaling pathways, including NK cell activation [36]. And the role of RhoA for phagocytosis has been studied in macrophages [37, 38]. Besides, RhoA signaling in T cells and B cells is pivotal for activation and migration [34]. The association of RRAP-mutant with immune activation and anti-tumor activity in the tumor immune microenvironment may provide clues to the predictive effect of RRAP in immunotherapy response.

The US Food and Drug Administration approved MSI-H/dMMR as a biomarker for immunotherapy; MSI-H/dMMR occurs in only 4-22% of gastric cancer cases [3, 39, 40] with an approximately 40-57% response [40, 41]. Many studies have shown that TMB, CD8+ TILs, and PD-L1 expression correlate with the efficacy of immunotherapy [42-44]; however, the association between these markers and the clinical benefit of gastric cancer immunotherapy is uncertain [45]. Accordingly, in the current study, 2 retrospective analyses on the IM1 and IM2 cohorts were performed to predict the efficacy of ICB therapy based on RRAP status. The response rate of RRAP-mutant patients was higher than that of RRAP-wild type patients within both the IM1 and IM2 cohorts. The differences are expected to be significant in a larger sized cohort. A clinical trial based on the RRAP biomarker is being designed. Our findings indicate that RRAP plays an important role in the regulation of the TME. Its status as a predictive biomarker will be further verified in larger clinical cohorts receiving immunotherapy.

In summary, this is one of the few efforts of biomarker identification for disease prognosis and therapeutic response based on pathway genomic characteristics in gastric cancer. In gastric cancer patients, RRAP-mutant tumors were correlated with a better prognosis, regardless of the histological classification and clinicopathological parameters. We investigated this correlation mainly from the perspective of RRAP regulating tumor invasion, metastasis, and the TME. Patients with RRAP-mutant tumors showed a better response to checkpoint inhibitor-based immunotherapy. These findings shed light on the clinical implications of the RRAP-mutant biomarker and may be used to guide personalized therapy for gastric cancer patients.

## MATERIALS AND METHODS

### Gastric cancer cohorts

The 4 gastric cancer patient cohorts included in this study were the TCGA cohort (N=315, Supplementary Table 3), the CGC cohort (N = 109, Supplementary Table 3), and 2 immunotherapy cohorts, IM1 (N = 37) and IM2 (N = 47). Clinical and mutation data from the

TCGA database (N = 440) were downloaded from the cBioportal database [46], and survival data were collected from a previous study [47]. Only patients with American Joint Committee on Cancer (AJCC) version 8 stage II and III were included (N = 319). Four patients without a survival time were excluded; the final number of patients in the TCGA cohort was 315. The TCGA training (N = 160) and validation (N = 155) cohorts were generated by random sampling to be approximately equal in sample size. Formalin-fixed, paraffin-embedded tumor tissue samples from the CGC cohort were collected from 109 stage II–III treatment-naive gastric cancer patients who had primary gastric cancer resection at Peking University Cancer Hospital and Institute between 2008 and 2015. The tumor tissue samples of the IM1 cohort were obtained from metastatic gastric cancer patients before they started ICB treatment. After the samples were collected, these 37 patients received at least 1 cycle of any ICB therapy regardless of the agent's target (i.e., PD-1/PD-L1). The IM2 cohort containing 47 patients who received PD-1 inhibitor therapy (toripalimab) was collected from multiple centers, details of which were previously published [48].

This study was approved by the medical ethics committee of the Peking University Cancer Hospital, and participants provided informed consent. The design and implementation of the study complied with the local regulations and guidelines and with the basic principles of the Declaration of Helsinki.

### **Whole-exome sequencing analysis and variant filtering**

Tumor tissue and adjacent nontumor tissue samples from CGC cohort patients were subjected to whole-exome sequencing. All whole-genome sequencing (including DNA extraction and quality control) was performed in the Origimed laboratory. This laboratory is College of American Pathologists-accredited and is a Clinical Laboratory Improvement Amendments (CLIA) certified laboratory. All tumor tissue slides were reviewed by two independent pathologists, and samples with estimated tumor purity greater than 20% were included in the study. In detail, DNA was extracted from the formalin-fixed, paraffin-embedded samples according to the manufacturer's instructions. Next, ~ 500 ng of genomic DNA was sheared to a mean fragment length of 200 bp and labeled with a 6–8 base barcode during polymerase chain reaction (PCR) amplification. Exomes were captured using a SureSelectXT Human All Exon V6 (Agilent Technologies). Sequencing was performed with an Illumina HiSeqX instrument using 150 base paired-end reads. Reads were trimmed with AdapterRemoval v2

[49], aligned to the human reference genome (GRCh37) by Burrows-Wheeler Aligner v0.7.5a [50], and PCR duplicates were removed by Picard (version 1.47). Somatic variants were identified by GATK4 Mutect2 (version 4.0.6) [51] and then annotated with SnpEff (version 4.3b) [52]. Variants in the common dbSNP database (version 147) or those having a frequency above 1.5% in the Exome Sequencing Project 6500 or 1000 genome project were excluded from further consideration. The variant files have been uploaded to the European Variation Archive (PRJEB31906). Low-confidence variants in both the TCGA and CGC cohorts were removed by applying the following filters: (1) total coverage < 30, (2) variant allele depth < 7, and (3) variant allele frequency < 0.05.

### **Feature selection and biomarker validation**

The prognostic biomarker was calculated by a genetic algorithm that was implemented in the Python package *pyeasyga* [50, 53]. In brief, each solution was a binary vector with a 0 or 1 value that had the same length as the original full set of the “regulation of RhoA activity pathway” biomarker (length = 48); 1 indicated that the corresponding gene harbored a nonsynonymous substitution or indel in the coding region, and 0 indicated neither of these changes. The goal was to find the best solution containing the genes with their values equal to 1, which was considered as the optimal gene subset (i.e., biomarker). The parameters of the genetic algorithm were empirically set as follows: a total population size of 40 solutions, crossover probability of 0.1, mutation probability of 0.01, and max generation of 50 000. The fitness value was calculated as follows: a cohort that was randomly selected (size = 80, 50%) from the 160 patients of the TCGA training cohort was further divided into a mutation and wild-type group according to the mutation status of the genes that corresponded to value 1 of each solution; a log-rank p-value was calculated for this division, the whole process was performed 100 times, and the mean p-values were calculated as the fitness value. Finally, the best solution was generated, and the resulting genes with a value of 1 were regarded as the final subset (i.e., the biomarker).

### **Tumor mutational burden, neoantigens, and mutational signatures**

The tumor mutational burden (TMB) was estimated by dividing the total number of somatic variants by the coding region size. All somatic variants in the coding and splicing regions were counted, and the coding region size was estimated as 33 Mb based on RefSeq Genome Annotation (GRCh37). HLA typing was performed with OptiType (version 1.2.1) [54]. All nonsynonymous mutations that were identified were

translated into peptides of 17 amino acids with an in-house pipeline. The sliding window (approximately 9-11 amino acids in size) method was used to identify substrings within the 17mer that had a predicted HLA class I binding affinity of less than 500 nM to any patient-specific HLA allele. The binding affinity for amino acids and alleles was analyzed using NetMHCpan v3.0 [55]. The neoantigen burden was estimated as the total number of substrings with a binding affinity less than 500 nM. Mutational signature contributions were identified using the R package deconstrucSigs (version 1.8.0) [56] with 30 signatures documented by the Catalogue of Somatic Mutations in Cancer as a reference [57]; samples with an error rate less than 0.15 and mutation counts greater than 30 were considered for mutational signature analysis. Deficient mismatch repair was estimated as the sum of signature 6, 15, 20, and 26 [57].

### RNA-seq data analysis

RNA-seq data of the TCGA cohort were extracted with TCGAAbiolinks [58]. Raw read counts were normalized with DESeq2 [59] and then log<sub>2</sub> transformed. The ssGSEA algorithm in the R package GSVA [60] was used to evaluate the pathway enrichment score. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway gene sets were retrieved from the KEGG database by the R package gage [61].

### Immunohistochemistry staining and evaluation

Immunohistochemistry (IHC) analysis for CD8 (clone SP16, ZSGB-BIO) was evaluated within intratumor areas. Aperio Scanscope (Aperio Technologies Vista, CA, USA) was used to quantify CD8<sup>+</sup> density by the rare event tissue test method. We counted the total number of CD8<sup>+</sup> cells of each area based on 6 randomly captured visual fields (400×400 μm<sup>2</sup>) and defined the density of CD8<sup>+</sup> TILs as the total cell number per square millimeter. IHC staining of anti-PD-L1 (clone SP142, Spring Bioscience) was also annotated within intratumor areas. The percentages of cancer cells and immune cells with anti-PD-L1 staining were measured in each area based on 3 visual fields in darkly stained areas (400×400 μm<sup>2</sup>). The expression of PD-L1 was defined as positive when ≥ 1% of the tumor/stromal cells were positive. IHC-stained sections were scored independently by 2 gastrointestinal pathologists blinded to the clinicopathological parameters and biomarker results.

### Mismatch repair/microsatellite instability testing

To evaluate the mismatch repair (MMR) and/or microsatellite instability (MSI) status, MLH1 (clone

ES05, Gene Tech), MSH2 (clone 25D12, Gene Tech), MSH6 (clone EP49, Gene Tech) and PMS2 (clone EP51, Gene Tech) were stained. The complete loss of expression of one or more proteins was considered as dMMR. In some cases, the microsatellite stability status was calculated by using a single multiplex PCR that assessed five microsatellite loci (BAT-25, BAT-26, D2S123, D5S346, and D17S250) [62]. For interpretation, instability at more than one locus referred to high microsatellite instability (MSI-H), instability at a single locus referred to low microsatellite instability (MSI-L), and no instability at any locus referred to stable microsatellites (MSS) [63].

### Statistical analysis

Survival analysis was carried out by the Kaplan-Meier method, and the difference between the groups was compared using the log-rank test. The hazard ratio and the 95% confidence interval were estimated by the Cox model. The multivariate Cox model was performed to adjust for confounding factors. Fisher's exact test was used to compare proportions. The Wilcoxon rank-sum test was used to compare continuous values. All statistical analyses were performed with R software 3.5.3.

### Abbreviations

CGC: Chinese gastric cancer; HLA: human leukocyte antigen; ICB: immune checkpoint blockade; IHC: immunohistochemistry; MMR: mismatch repair; MSI: microsatellite instability; MSS: microsatellite stable; NAB: neoantigen burden; OS: overall survival; PFS: progression free survival; RRAP: regulation of RhoA activity panel; TCGA: The Cancer Genome Atlas; TILs: infiltrating lymphocytes; TMB: tumor mutational burden; TME: tumor microenvironment.

### AUTHOR CONTRIBUTIONS

LS, WS and JG conceived and designed the study. WH and SZ performed the experiments and analyzed the data. CZ, ZL and SG performed the sample preparation and the pathological evaluation. BL, HF, and KW contributed the reagents, materials, and analysis tools. RX and JJ involved in the sample collection and provided the diagnostic data. WH, SZ and WS wrote the manuscript. All authors read and approved the final manuscript.

### ACKNOWLEDGMENTS

We thank all the patients who participated in this study, their families; the clinicians and the staff of Peking University Cancer Hospital and Institute, for data management and logistic support.

## CONFLICTS OF INTEREST

The authors have declared that no conflicts of interest exists.

## FUNDING

This research was supported by funding from the National Natural Science Foundation of China (No. 91959205, and No. 81872341), and the National Key Research and Development Program of China (No. 2017YFC1308900).

## REFERENCES

1. Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, Brenner H, Dicker DJ, Chimed-Orchir O, Dandona R, Dandona L, Fleming T, Forouzanfar MH, Hancock J, et al, and Global Burden of Disease Cancer Collaboration. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study. *JAMA Oncol.* 2017; 3:524–48.  
<https://doi.org/10.1001/jamaoncol.2016.5688>  
PMID:27918777
2. Lauren P. The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. An attempt at a histo-clinical classification. *Acta Pathol Microbiol Scand.* 1965; 64:31–49.  
<https://doi.org/10.1111/apm.1965.64.1.31>  
PMID:14320675
3. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature.* 2014; 513:202–09.  
<https://doi.org/10.1038/nature13480>  
PMID:25079317
4. Cristescu R, Lee J, Nebozhyn M, Kim KM, Ting JC, Wong SS, Liu J, Yue YG, Wang J, Yu K, Ye XS, Do IG, Liu S, et al. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nat Med.* 2015; 21:449–56.  
<https://doi.org/10.1038/nm.3850>  
PMID:25894828
5. Mitchison TJ, Cramer LP. Actin-based cell motility and cell locomotion. *Cell.* 1996; 84:371–79.  
[https://doi.org/10.1016/s0092-8674\(00\)81281-7](https://doi.org/10.1016/s0092-8674(00)81281-7)  
PMID:8608590
6. Lauffenburger DA, Horwitz AF. Cell migration: a physically integrated molecular process. *Cell.* 1996; 84:359–69.  
[https://doi.org/10.1016/s0092-8674\(00\)81280-5](https://doi.org/10.1016/s0092-8674(00)81280-5)  
PMID:8608589
7. Olson MF, Paterson HF, Marshall CJ. Signals from ras and rho GTPases interact to regulate expression of p21Waf1/Cip1. *Nature.* 1998; 394:295–99.  
<https://doi.org/10.1038/28425> PMID:9685162
8. Kalpana G, Figy C, Yeung M, Yeung KC. Reduced RhoA expression enhances breast cancer metastasis with a concomitant increase in CCR5 and CXCR4 chemokines signaling. *Sci Rep.* 2019; 9:16351.  
<https://doi.org/10.1038/s41598-019-52746-w>  
PMID:31705019
9. Nam GH, Lee EJ, Kim YK, Hong Y, Choi Y, Ryu MJ, Woo J, Cho Y, Ahn DJ, Yang Y, Kwon IC, Park SY, Kim IS. Combined rho-kinase inhibition and immunogenic cell death triggers and propagates immunity against cancer. *Nat Commun.* 2018; 9:2165.  
<https://doi.org/10.1038/s41467-018-04607-9>  
PMID:29867097
10. Huang KH, Lan YT, Chen MH, Chao Y, Lo SS, Li AF, Wu CW, Chiou SH, Yang MH, Shyr YM, Fang WL. The correlation between RhoA expression and clinicopathological characteristics in gastric cancer patients after curative surgery. *World J Surg.* 2015; 39:2289–99.  
<https://doi.org/10.1007/s00268-015-3095-4>  
PMID:26013205
11. Ushiku T, Ishikawa S, Kakiuchi M, Tanaka A, Katoh H, Aburatani H, Lauwers GY, Fukayama M. RHOA mutation in diffuse-type gastric cancer: a comparative clinicopathology analysis of 87 cases. *Gastric Cancer.* 2016; 19:403–11.  
<https://doi.org/10.1007/s10120-015-0493-0>  
PMID:25823974
12. Röcken C, Behrens HM, Böger C, Krüger S. Clinicopathological characteristics of RHOA mutations in a Central European gastric cancer cohort. *J Clin Pathol.* 2016; 69:70–75.  
<https://doi.org/10.1136/jclinpath-2015-202980>  
PMID:26251521
13. Shi W, Jiang T, Nuciforo P, Hatzis C, Holmes E, Harbeck N, Sotiriou C, Peña L, Loi S, Rosa DD, Chia S, Wardley A, Ueno T, et al. Pathway level alterations rather than mutations in single genes predict response to HER2-targeted therapies in the neo-ALTTO trial. *Ann Oncol.* 2017; 28:128–35.  
<https://doi.org/10.1093/annonc/mdw434>  
PMID:28177460
14. Sahai E, Marshall CJ. RHO-GTPases and cancer. *Nat Rev Cancer.* 2002; 2:133–42.  
<https://doi.org/10.1038/nrc725> PMID:12635176
15. Wang K, Yuen ST, Xu J, Lee SP, Yan HH, Shi ST, Siu HC, Deng S, Chu KM, Law S, Chan KH, Chan AS, Tsui WY, et al. Whole-genome sequencing and comprehensive



- molecular profiling identify new driver mutations in gastric cancer. *Nat Genet.* 2014; 46:573–82.  
<https://doi.org/10.1038/ng.2983>  
PMID:24816253
16. Kakiuchi M, Nishizawa T, Ueda H, Gotoh K, Tanaka A, Hayashi A, Yamamoto S, Tatsuno K, Katoh H, Watanabe Y, Ichimura T, Ushiku T, Funahashi S, et al. Recurrent gain-of-function mutations of RHOA in diffuse-type gastric carcinoma. *Nat Genet.* 2014; 46:583–87.  
<https://doi.org/10.1038/ng.2984> PMID:24816255
  17. Xu XT, Song QB, Yao Y, Ruan P, Tao ZZ. Inhibition of RhoA/ROCK signaling pathway promotes the apoptosis of gastric cancer cells. *Hepatogastroenterology.* 2012; 59:2523–26.  
<https://doi.org/10.5754/hge12147> PMID:22584506
  18. Murray D, Horgan G, Macmathuna P, Doran P. NET1-mediated RhoA activation facilitates lysophosphatidic acid-induced cell migration and invasion in gastric cancer. *Br J Cancer.* 2008; 99:1322–29.  
<https://doi.org/10.1038/sj.bjc.6604688>  
PMID:18827818
  19. Leyden J, Murray D, Moss A, Arumuguma M, Doyle E, McEntee G, O’Keane C, Doran P, MacMathuna P. Net1 and myeov: computationally identified mediators of gastric cancer. *Br J Cancer.* 2006; 94:1204–12.  
<https://doi.org/10.1038/sj.bjc.6603054>  
PMID:16552434
  20. Lin KY, Wang LH, Hseu YC, Fang CL, Yang HL, Kumar KJ, Tai C, Uen YH. Clinical significance of increased guanine nucleotide exchange factor Vav3 expression in human gastric cancer. *Mol Cancer Res.* 2012; 10:750–59.  
<https://doi.org/10.1158/1541-7786.MCR-11-0598-T>  
PMID:22544459
  21. Zhang H, Schaefer A, Wang Y, Hodge RG, Blake DR, Diehl JN, Papageorge AG, Stachler MD, Liao J, Zhou J, Wu Z, Akarca FG, de Klerk LK, et al. Gain-of-function RHOA mutations promote focal adhesion kinase activation and dependency in diffuse gastric cancer. *Cancer Discov.* 2020; 10:288–305.  
<https://doi.org/10.1158/2159-8290.CD-19-0811>  
PMID:31771969
  22. Havel JJ, Chowell D, Chan TA. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. *Nat Rev Cancer.* 2019; 19:133–50.  
<https://doi.org/10.1038/s41568-019-0116-x>  
PMID:30755690
  23. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, Ziv E, Culhane AC, Paull EO, et al. The Immune Landscape of Cancer. *Immunity.* 2018; 48:812–830.e14.  
<https://doi.org/10.1016/j.immuni.2018.03.023>  
PMID:29628290
  24. Pagès F, Berger A, Camus M, Sanchez-Cabo F, Costes A, Molitor R, Mlecnik B, Kirilovsky A, Nilsson M, Damotte D, Meatchi T, Bruneval P, Cugnenc PH, et al. Effector memory T cells, early metastasis, and survival in colorectal cancer. *N Engl J Med.* 2005; 353:2654–66.  
<https://doi.org/10.1056/NEJMoa051424>  
PMID:16371631
  25. Ohri CM, Shikotra A, Green RH, Waller DA, Bradding P. Macrophages within NSCLC tumour islets are predominantly of a cytotoxic M1 phenotype associated with extended survival. *Eur Respir J.* 2009; 33:118–26.  
<https://doi.org/10.1183/09031936.00065708>  
PMID:19118225
  26. Martínez-Lostao L, Anel A, Pardo J. How do cytotoxic lymphocytes kill cancer cells? *Clin Cancer Res.* 2015; 21:5047–56.  
<https://doi.org/10.1158/1078-0432.CCR-15-0685>  
PMID:26567364
  27. Trapani JA, Smyth MJ. Functional significance of the perforin/granzyme cell death pathway. *Nat Rev Immunol.* 2002; 2:735–47.  
<https://doi.org/10.1038/nri911> PMID:12360212
  28. Green DR, Ferguson TA. The role of fas ligand in immune privilege. *Nat Rev Mol Cell Biol.* 2001; 2:917–24.  
<https://doi.org/10.1038/35103104>  
PMID:11733771
  29. Narumiya S, Tanji M, Ishizaki T. Rho signaling, ROCK and mDia1, in transformation, metastasis and invasion. *Cancer Metastasis Rev.* 2009; 28:65–76.  
<https://doi.org/10.1007/s10555-008-9170-7>  
PMID:19160018
  30. Yoon C, Cho SJ, Aksoy BA, Park DJ, Schultz N, Ryeom SW, Yoon SS. Chemotherapy resistance in diffuse-type gastric adenocarcinoma is mediated by RhoA activation in cancer stem-like cells. *Clin Cancer Res.* 2016; 22:971–83.  
<https://doi.org/10.1158/1078-0432.CCR-15-1356>  
PMID:26482039
  31. Chang KK, Cho SJ, Yoon C, Lee JH, Park DJ, Yoon SS. Increased RhoA activity predicts worse overall survival in patients undergoing surgical resection for lauren diffuse-type gastric adenocarcinoma. *Ann Surg Oncol.* 2016; 23:4238–46.  
<https://doi.org/10.1245/s10434-016-5357-2>  
PMID:27364501
  32. Harding CV, Geuze HJ. Class II MHC molecules are present in macrophage lysosomes and phagolysosomes that function in the phagocytic processing of listeria monocytogenes for presentation to T cells. *J Cell Biol.* 1992; 119:531–42.  
<https://doi.org/10.1083/jcb.119.3.531>  
PMID:1400590

33. König R, Huang LY, Germain RN. MHC class II interaction with CD4 mediated by a region analogous to the MHC class I binding site for CD8. *Nature*. 1992; 356:796–98.  
<https://doi.org/10.1038/356796a0> PMID:1574118
34. Bros M, Haas K, Moll L, Grabbe S. RhoA as a key regulator of innate and adaptive immunity. *Cells*. 2019; 8:733.  
<https://doi.org/10.3390/cells8070733> PMID:31319592
35. Al Absi A, Wurzer H, Guerin C, Hoffmann C, Moreau F, Mao X, Brown-Clay J, Petrolli R, Casellas CP, Dieterle M, Thiery JP, Chouaib S, Berchem G, et al. Actin cytoskeleton remodeling drives breast cancer cell escape from natural killer-mediated cytotoxicity. *Cancer Res*. 2018; 78:5631–43.  
<https://doi.org/10.1158/0008-5472.CAN-18-0441> PMID:30104240
36. Rodriguez RM, Suarez-Alvarez B, Lavín JL, Mosén-Ansorena D, Baragaño Raneros A, Márquez-Kisinousky L, Aransay AM, Lopez-Larrea C. Epigenetic networks regulate the transcriptional program in memory and terminally differentiated CD8<sup>+</sup> T cells. *J Immunol*. 2017; 198:937–49.  
<https://doi.org/10.4049/jimmunol.1601102> PMID:27974453
37. Nakaya M, Tanaka M, Okabe Y, Hanayama R, Nagata S. Opposite effects of rho family GTPases on engulfment of apoptotic cells by macrophages. *J Biol Chem*. 2006; 281:8836–42.  
<https://doi.org/10.1074/jbc.M510972200> PMID:16439364
38. Kim SY, Kim S, Bae DJ, Park SY, Lee GY, Park GM, Kim IS. Coordinated balance of Rac1 and RhoA plays key roles in determining phagocytic appetite. *PLoS One*. 2017; 12:e0174603.  
<https://doi.org/10.1371/journal.pone.0174603> PMID:28376111
39. Salem ME, Puccini A, Grothey A, Raghavan D, Goldberg RM, Xiu J, Korn WM, Weinberg BA, Hwang JJ, Shields AF, Marshall JL, Philip PA, Lenz HJ. Landscape of tumor mutation load, mismatch repair deficiency, and PD-L1 expression in a large patient cohort of gastrointestinal cancers. *Mol Cancer Res*. 2018; 16:805–12.  
<https://doi.org/10.1158/1541-7786.MCR-17-0735> PMID:29523759
40. Fuchs CS, Doi T, Jang RW, Muro K, Satoh T, Machado M, Sun W, Jalal SI, Shah MA, Metges JP, Garrido M, Golan T, Mandala M, et al. Safety and efficacy of pembrolizumab monotherapy in patients with previously treated advanced gastric and gastroesophageal junction cancer: phase 2 clinical KEYNOTE-059 trial. *JAMA Oncol*. 2018; 4:e180013.  
<https://doi.org/10.1001/jamaoncol.2018.0013> PMID:29543932
41. Marabelle A, Le DT, Ascierto PA, Di Giacomo AM, De Jesus-Acosta A, Delord JP, Geva R, Gottfried M, Penel N, Hansen AR, Piha-Paul SA, Doi T, Gao B, et al. Efficacy of pembrolizumab in patients with noncolorectal high microsatellite instability/mismatch repair-deficient cancer: results from the phase II KEYNOTE-158 study. *J Clin Oncol*. 2020; 38:1–10.  
<https://doi.org/10.1200/JCO.19.02105> PMID:31682550
42. Zhu J, Powis de Tenbossche CG, Cané S, Colau D, van Baren N, Lurquin C, Schmitt-Verhulst AM, Liljeström P, Uyttenhove C, Van den Eynde BJ. Resistance to cancer immunotherapy mediated by apoptosis of tumor-infiltrating lymphocytes. *Nat Commun*. 2017; 8:1404.  
<https://doi.org/10.1038/s41467-017-00784-1> PMID:29123081
43. Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Csőszi T, Fülöp A, Gottfried M, Peled N, Tafreshi A, Cuffe S, O'Brien M, Rao S, Hotta K, et al, and KEYNOTE-024 Investigators. Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. *N Engl J Med*. 2016; 375:1823–33.  
<https://doi.org/10.1056/NEJMoa1606774> PMID:27718847
44. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, Lee W, Yuan J, Wong P, Ho TS, Miller ML, Rekhtman N, Moreira AL, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science*. 2015; 348:124–28.  
<https://doi.org/10.1126/science.aaa1348> PMID:25765070
45. Shitara K, Özgüroğlu M, Bang YJ, Di Bartolomeo M, Mandalà M, Ryu MH, Fornaro L, Olesiński T, Caglevic C, Chung HC, Muro K, Goekkurt E, Mansoor W, et al, and KEYNOTE-061 investigators. Pembrolizumab versus paclitaxel for previously treated, advanced gastric or gastro-oesophageal junction cancer (KEYNOTE-061): a randomised, open-label, controlled, phase 3 trial. *Lancet*. 2018; 392:123–33.  
[https://doi.org/10.1016/S0140-6736\(18\)31257-1](https://doi.org/10.1016/S0140-6736(18)31257-1) PMID:29880231
46. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. 2012; 2:401–04.  
<https://doi.org/10.1158/2159-8290.CD-12-0095> PMID:22588877
47. Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, Kovatich AJ, Benz CC, Levine DA, Lee AV,

- Omberg L, Wolf DM, Shriver CD, et al, and Cancer Genome Atlas Research Network. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell*. 2018; 173:400–16.e11.  
<https://doi.org/10.1016/j.cell.2018.02.052>  
PMID:29625055
48. Wang F, Wei XL, Wang FH, Xu N, Shen L, Dai GH, Yuan XL, Chen Y, Yang SJ, Shi JH, Hu XC, Lin XY, Zhang QY, et al. Safety, efficacy and tumor mutational burden as a biomarker of overall survival benefit in chemo-refractory gastric cancer treated with toripalimab, a PD-1 antibody in phase Ib/II clinical trial NCT02915432. *Ann Oncol*. 2019; 30:1479–86.  
<https://doi.org/10.1093/annonc/mdz197>  
PMID:31236579
49. Schubert M, Lindgreen S, Orlando L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res Notes*. 2016; 9:88.  
<https://doi.org/10.1186/s13104-016-1900-2>  
PMID:26868221
50. Gould Rothberg BE, Berger AJ, Molinaro AM, Subtil A, Krauthammer MO, Camp RL, Bradley WR, Ariyan S, Kluger HM, Rimm DL. Melanoma prognostic model using tissue microarrays and genetic algorithms. *J Clin Oncol*. 2009; 27:5772–80.  
<https://doi.org/10.1200/JCO.2009.22.8239>  
PMID:19884546
51. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol*. 2013; 31:213–19.  
<https://doi.org/10.1038/nbt.2514> PMID:23396013
52. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)*. 2012; 6:80–92.  
<https://doi.org/10.4161/fly.19695> PMID:22728672
53. Dolled-Filhart M, Rydén L, Cregger M, Jirström K, Harigopal M, Camp RL, Rimm DL. Classification of breast cancer using genetic algorithms and tissue microarrays. *Clin Cancer Res*. 2006; 12:6459–68.  
<https://doi.org/10.1158/1078-0432.CCR-06-1383>  
PMID:17085660
54. Szolek A, Schubert B, Mohr C, Sturm M, Feldhahn M, Kohlbacher O. OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics*. 2014; 30:3310–16.  
<https://doi.org/10.1093/bioinformatics/btu548>  
PMID:25143287
55. Nielsen M, Andreatta M. NetMHCpan-3.0; improved prediction of binding to MHC class I molecules integrating information from multiple receptor and peptide length datasets. *Genome Med*. 2016; 8:33.  
<https://doi.org/10.1186/s13073-016-0288-x>  
PMID:27029192
56. Rosenthal R, McGranahan N, Herrero J, Taylor BS, Swanton C. DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol*. 2016; 17:31.  
<https://doi.org/10.1186/s13059-016-0893-4>  
PMID:26899170
57. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Børresen-Dale AL, Boyault S, Burkhardt B, Butler AP, et al, and Australian Pancreatic Cancer Genome Initiative, and ICGC Breast Cancer Consortium, and ICGC MML-Seq Consortium, and ICGC PedBrain. Signatures of mutational processes in human cancer. *Nature*. 2013; 500:415–21.  
<https://doi.org/10.1038/nature12477>  
PMID:23945592
58. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Carolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I, Ceccarelli M, Bontempi G, Noushmehr H. TCGAbiolinks: an R/bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res*. 2016; 44:e71.  
<https://doi.org/10.1093/nar/gkv1507>  
PMID:26704973
59. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014; 15:550.  
<https://doi.org/10.1186/s13059-014-0550-8>  
PMID:25516281
60. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013; 14:7.  
<https://doi.org/10.1186/1471-2105-14-7>  
PMID:23323831
61. Luo W, Friedman MS, Shedden K, Hankenson KD, Woolf PJ. GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics*. 2009; 10:161.  
<https://doi.org/10.1186/1471-2105-10-161>  
PMID:19473525
62. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, Meltzer SJ, Rodriguez-Bigas MA, Fodde R, Ranzani GN, Srivastava S. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the

determination of microsatellite instability in colorectal cancer. *Cancer Res.* 1998; 58:5248–57.  
PMID:[9823339](https://pubmed.ncbi.nlm.nih.gov/9823339/)

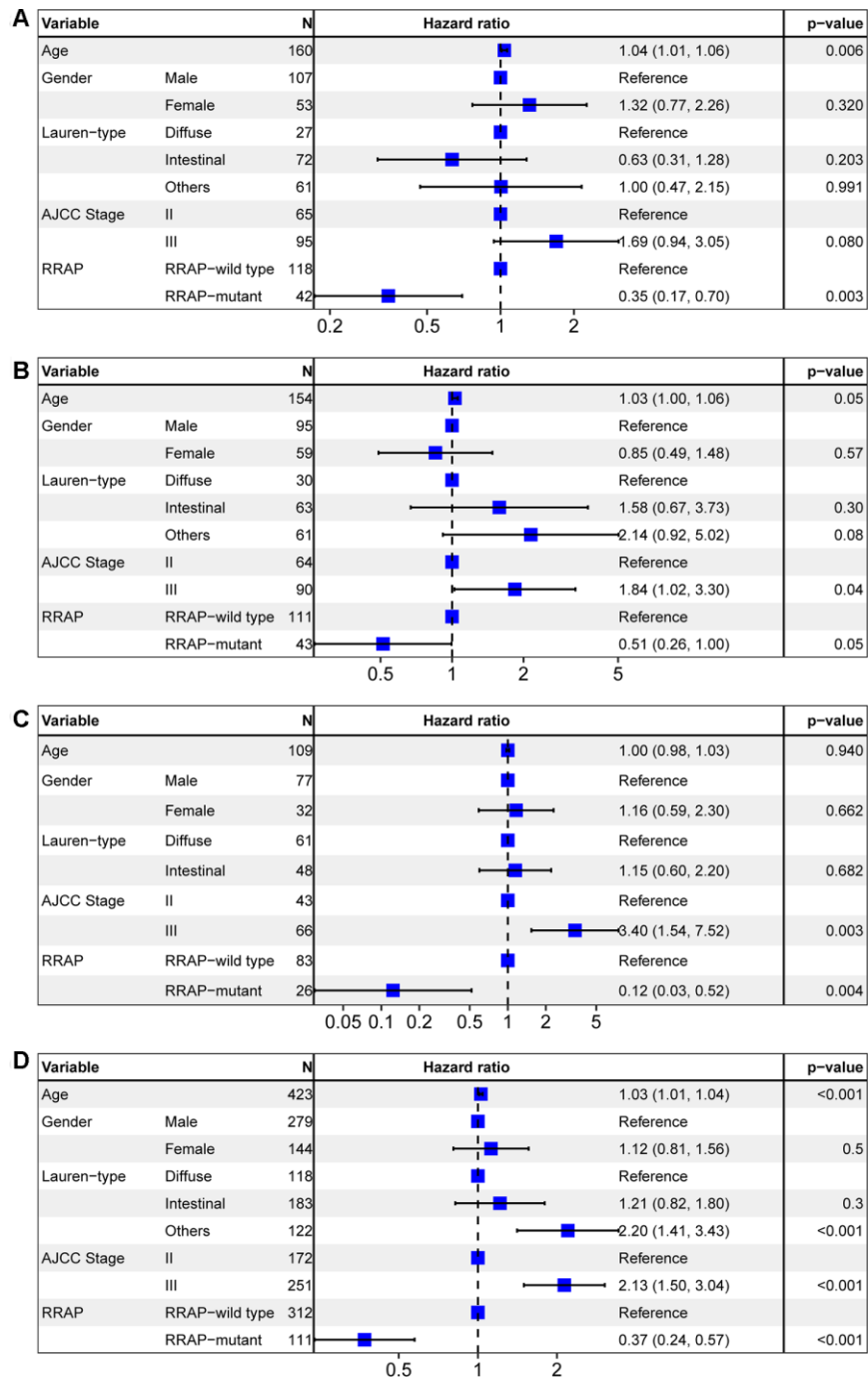
63. Jang M, Kwon Y, Kim H, Kim H, Min BS, Park Y, Kim TI, Hong SP, Kim WK. Microsatellite instability test using

peptide nucleic acid probe-mediated melting point analysis: a comparison study. *BMC Cancer.* 2018; 18:1218.  
<https://doi.org/10.1186/s12885-018-5127-6>  
PMID:[30514259](https://pubmed.ncbi.nlm.nih.gov/30514259/)

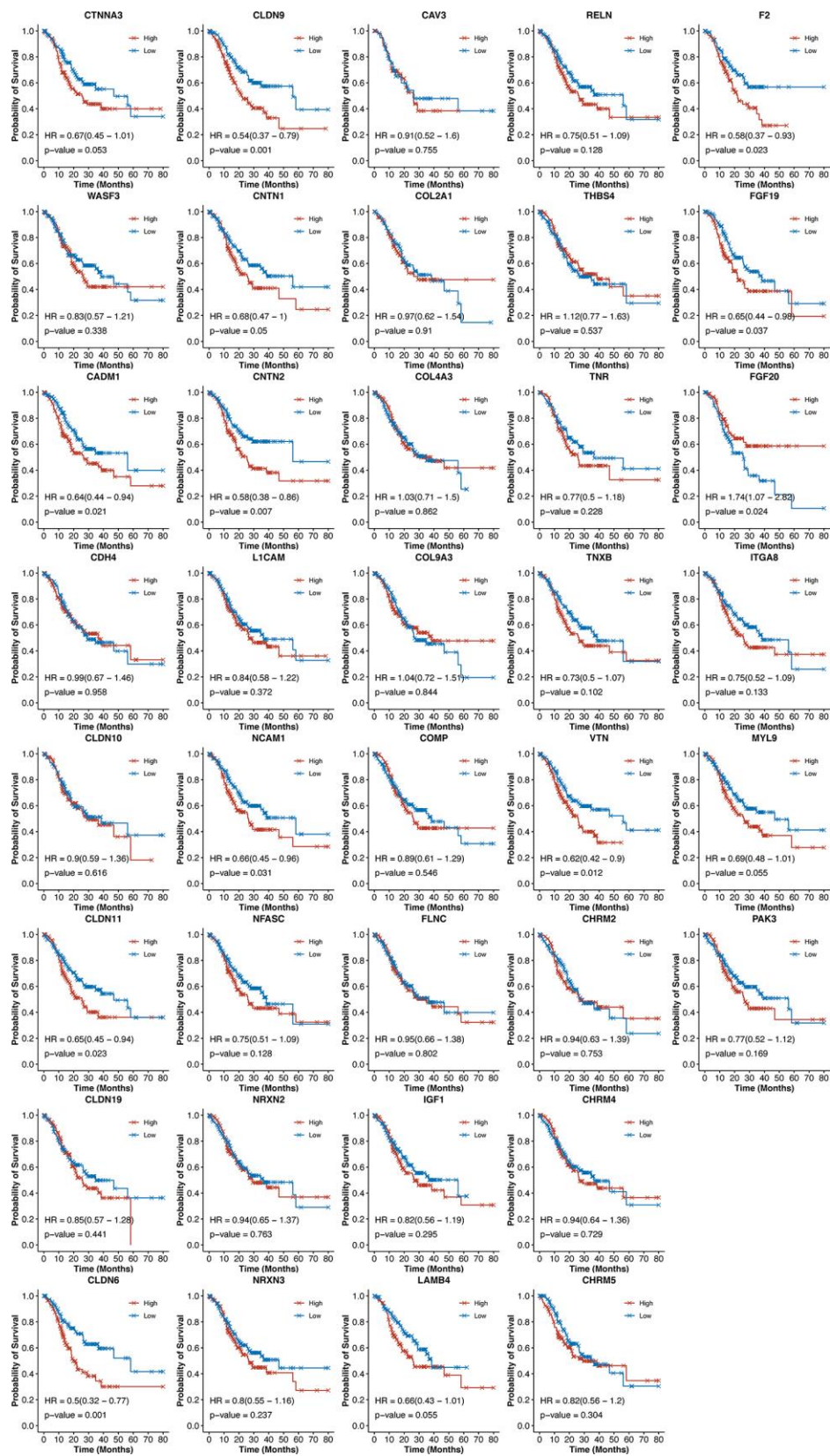


SUPPLEMENTARY MATERIALS

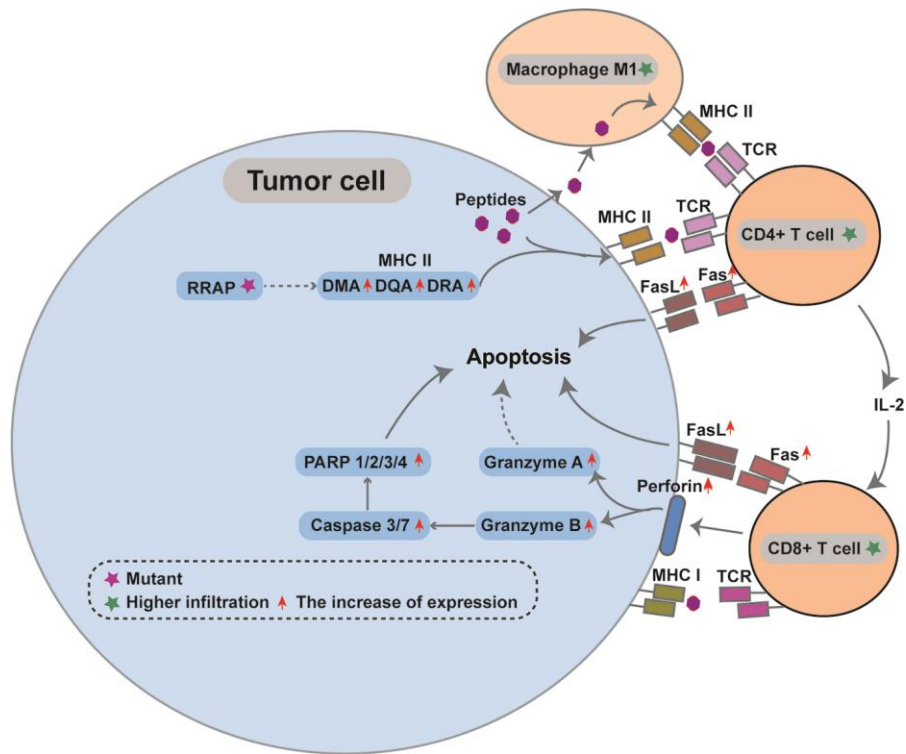
Supplementary Figures



**Supplementary Figure 1.** Multivariate Cox model analysis, with age, gender, Lauren type, AJCC stage, and RRAP as confounding factors, performed on the TCGA training cohort (A), TCGA validation cohort (B), CGC cohort (C), and the overall cohort containing merged data from the three cohorts (D).



Supplementary Figure 2. Kaplan-Meier overall survival curves for patients with high and low expression of 39 differentially expressed genes of the 4 migration-related functions. P-values were estimated with the log-rank test, and hazard ratios were estimated with the Cox model.



Supplementary Figure 3. Diagram of the interaction of immune cells and tumor cells in the RRAP-mutant tumor microenvironment.

## Supplementary Tables

Please browse Full Text version to see the data of Supplementary Table 1.

**Supplementary Table 1. Genes contained in the RRAP biomarker gene set and their mutation frequency in the TCGA training, TCGA validation, and CGC cohorts, along with migration activity.**

**Supplementary Table 2. The CLDN11, CLDN6, CLDN9, VTN, and F2 with cell migration activity.**

Symbol	Conclusion	Experiment model	PMID
CLDN6	Higher <b>CLDN6</b> expression promotes the cell proliferation and migration ability of gastric cancer. Increased expression of <b>CLDN6</b> predicts poor prognosis in gastric cancer patients.	MKN28, AGS, MKN7, NUGC-3, xenograft subcutaneous tumor model	31827075, 20874001, 27914788, 31654186
CLDN9	<b>CLDN9</b> overexpression in gastric adenocarcinoma cells has been reported to increase cell migration and proliferation. The expression of <b>CLDN9</b> in gastric cancer correlates to poor prognosis. Higher <b>CLDN9</b> expression promotes invasiveness of several solid tumors in vitro and metastasis in vivo.	AGS, HL7702, a spontaneous metastasis model	20874001, 29031421, 31418417, 26669782
CLDN11	Silencing of <b>CLDN11</b> is associated with increased invasiveness, proliferation, and migration of gastric cancer cells. DNA hypermethylation is associated with the downregulation of <b>CLDN11</b> in gastric cancer cells.	MKN28, MKN74, AGS, HFE145	32119960, 28962204, 19956721
F2	<b>F2</b> is generated in the TME, promoting the migration and metastasis of tumor cells.	B16, HT29-D4, SUM149, MC38 cells	30654498, 15539922, 12707033, 23280128
VTN	Decreased expression of <b>VTN</b> promotes gastric cancer cell growth and metastasis. <b>VTN</b> is related to the migration and invasion of other solid tumors, such as ovarian adenocarcinoma, fibrosarcoma and nasopharyngeal carcinoma.	AGS, IGROV1, HT-1080,5-8F, HNE2, HONE1 cells	25789040, 9935212, 16052409, 29123267



**Supplementary Table 3. Clinical characteristics in TCGA-training, TCGA-validation, and CGC cohorts.<sup>a</sup>**

<b>Characteristic</b>	<b>TCGA- training (N=160)</b>	<b>TCGA-validation (N=155)</b>	<b>CGC (N=109)</b>
Age, years			
Mean (SD)	64.3 (10.8)	66.2 (9.9)	57.9 (13.0)
Median (Q1 – Q3)	66 (57-72)	68 (58-73)	58 (51-67)
Min – Max	30-84	43-90	25-83
Gender - no. (%)			
Male	107 (66.9%)	96 (61.9%)	77 (70.6%)
Female	53 (33.1%)	59 (38.1%)	32 (29.4%)
AJCC Stage - no. (%)			
II	65 (40.6%)	65 (41.9%)	43 (39.4%)
III	95 (59.4%)	90 (58.1%)	66 (60.6%)
Lauren classification - no. (%)			
Diffuse	27 (16.9%)	30 (19.4%)	61 (56.0%)
Intestinal	72 (45.0%)	64 (41.3%)	48 (44.0%)
Unknown	61 (38.1%)	61 (39.3%)	—
Tumor location - no. (%)			
Gastroesophageal junction	33 (21.0%)	31 (20.0%)	34 (31.2%)
Non-gastroesophageal junction	123 (76.9%)	121 (78.1%)	75 (68.8%)
Unknown	4 (3.0%)	3 (2.0%)	—
Adjuvant chemotherapy - no. (%)			
Yes	81 (50.6%)	70 (45.2%)	98 (90.0%)
No	—	—	8 (7.3%)
Unknown	79 (49.4%)	85 (54.8%)	3 (2.7%)

<sup>a</sup>Because of rounding, percentages may not total 100.