

## Elevated Ras related GTP binding B (RRAGB) expression predicts poor overall survival and constructs a prognostic nomogram for colon adenocarcinoma

Jianjia Xiao<sup>a,\*</sup>, Qingqing Liu<sup>b,\*</sup>, Weijie Wu<sup>c,\*</sup>, Ying Yuan<sup>d</sup>, Jie Zhou<sup>a</sup>, Jieyu Shi<sup>e</sup>, and Shaorong Zhou<sup>a</sup>

<sup>a</sup>Department of General Surgery, Taizhou Hospital of Traditional Chinese Medicine, Taizhou, Jiangsu Province, China; <sup>b</sup>Department of Gastroenterology, Affiliated Hospital NO.2 Of Nantong University, Nantong, Jiangsu Province, China; <sup>c</sup>Department of Orthopedics, The Sixth People's Hospital of Nantong, Medical College of Nantong University, Nantong, Jiangsu Province, China; <sup>d</sup>Department of Geriatrics, Taizhou Second People's Hospital, Taizhou, Jiangsu Province, China; <sup>e</sup>Department of Neurology, Taizhou Hospital of Traditional Chinese Medicine, Taizhou, Jiangsu Province, China

### ABSTRACT

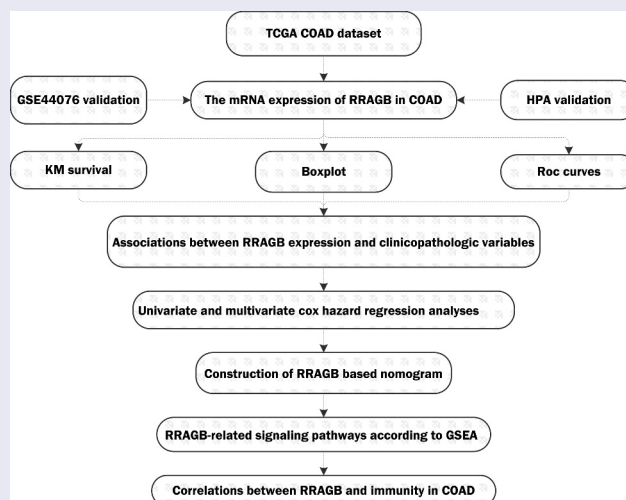
Currently, no articles have explored the roles of RRAGB gene in the occurrence and development of cancer. By means of The Cancer Genome Atlas (TCGA) data mining, we found that this gene might be a novel prognostic predictor for colon adenocarcinoma (COAD). Hence, this article was carried out to explore its roles in COAD and associations with immunity. RRAGB single-gene expression matrix and corresponding clinical information were extracted from TCGA database. Univariate/multivariate cox regression analyses and gene set enrichment analysis (GSEA) were utilized to identify independent prognostic factors and RRAGB related pathways, respectively. Relationships between RRAGB and immunity were also analyzed. Boxplot and K-M survival analysis indicated that RRAGB was not only differently expressed in COAD ( $P < 0.05$ ), but also significantly associated with overall survival (OS;  $P < 0.05$ ). Univariate and multivariate Cox hazard regression analyses indicated that RRAGB could serve as an independent prognostic factor for COAD (both  $P < 0.05$ ). GSEA identified five signaling pathways significantly enriched in the high-RRAGB expression phenotype. Moreover, a RRAGB-based nomogram was successfully constructed and displayed a satisfactory performance. In addition, RRAGB expression was found to be significantly associated with microsatellite instability (MSI), tumor mutational burden (TMB) and immunity. Our results revealed that RRAGB could be a prognostic biomarker for COAD in terms of OS and markedly related to MSI, TMB, and immunity. We also constructed an RRAGB-based nomogram with a satisfactory performance. Further researches should be carried out to validate our findings.



### ARTICLE HISTORY

Received 14 April 2021  
Revised 12 July 2021  
Accepted 13 July 2021


### KEYWORDS

Colon adenocarcinoma;  
RRAGB; nomogram; survival;  
TCGA



**CONTACT** Shaorong Zhou  [zhousr2005@126.com](mailto:zhousr2005@126.com)  Department of General Surgery, Taizhou Hospital of Traditional Chinese Medicine, No.86 Jichuan East Road, Taizhou, Jiangsu Province 225300, China

\*These three authors are contributed equally to this work.

 Supplemental data for this article can be accessed [here](#).

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Introduction

Colorectal cancer (CRC), as one of the most common tumor diagnosed in the digestive system, accounts for 8% with approximately 149,500 newly diagnosed cases and accounts for 9% with approximately 52,980 newly estimated death in the United States, 2021 [1]. Colon adenocarcinoma (COAD), as the most common histological subtype of colon cancer, mainly occurs in the intestinal mucosa and spreads to adjacent organs [2]. Although the 5-year survival rate for early-stage colon cancer patients undergoing radical resection is more than 90%, most patients are diagnosed with advanced cases or metastasis, leading to the drop of 5-year survival rates to 10% [3]. Currently, surgery, radiotherapy, chemotherapy, targeted therapies are available for the clinical treatment of COAD and considerable advancements have been achieved in these therapeutic strategies [4]. However, patients' prognosis is still poor and far from satisfactory due to the late diagnosis, rapid development and high frequency of metastasis [5,6]. Hence, there is an urgent need to explore the molecular mechanisms of COAD and to identify novel biomarkers for survival evaluation and targeted treatment [7].

Ras-related GTP binding B (RRAGB), also known as RAGB or bA465E19.1, belongs to the large family of Ras-homologous GTPases and encodes proteins of cellular switches operated by GTP-exchange factors and factors stimulating their intrinsic GTPase activity [8,9]. Thanks to the great improvement of high-throughput sequencing technology and an increasing establishment of public database networks, more and more transcriptome and clinical data could be explored by means of bioinformatics analysis. Based on these, a growing number of prognostic biomarkers or signatures had been established in different cancers [10–12]. As for RRAGB, Shi et al. successfully constructed a signature based on CHMP4C, FOXO1, RRAGB to effectively predict the cervical cancer patients' prognosis [13] and Xie et al. also constructed a six-gene model including RRAGB predict the non-small-cell lung cancer patients' overall survival (OS) [14]. Currently, little was still known about the roles of RRAGB gene in the occurrence and development of COAD. In this article, we aimed to explore the comprehensive

roles of RRAGB in COAD and to underline its associations with immunity, hoping to provide a novel candidate gene to improve the prognosis and survival rates of COAD.

## Materials and methods

### *Single gene matrix mining from the TCGA database*

RNA-sequencing FPKM data and corresponding clinical information of 39 normal and 398 COAD tumor tissues were extracted from the official website of The Cancer Genome Atlas (TCGA; <https://tcga-data.nci.nih.gov/tcga/>). R version 3.5.1 software (<https://www.r-project.org/>) was utilized to standardize RRAGB RNA-Sequencing data [15] and we further did an overlap with RRAGB mRNA to get single-gene expression matrix and clinical information for each sample ID. OS was the primary outcome of this study and we further analyzed the associations between RRAGB's expression and clinical data. Besides, we also used 'limma' package (<http://www.bioconductor.org/packages/release/bioc/html/limma.html>) to calculate differently expressed genes (DEGs), with the cutoff criteria of adjusted P-value (FDR) <0.05 and  $|\log_2 \text{FC (fold change)}| \geq 1$  [16].

### *Gene set enrichment analysis (GSEA)*

As a reference, the 'c2.cp.kegg.v6.2.symbols.gmt' gene set was obtained from the Molecular Signatures Database (MSigDB) (<http://software.broadinstitute.org/gsea/msigdb>) [17]. We performed GSEA to reveal significant survival differences between the high-RRAGB and low-RRAGB expression groups, with at least 1000 times permutation tests for each analysis [18]. Through this way we could discover significant critical biological pathways, with the threshold of the normalized enrichment score (NES)>1.5 and normal p values <0.05.

### *Protein–protein interaction (PPI) network and the Human Protein Atlas (HPA) database*

The PPI network analysis was carried out with the help of the online STRING (<https://string-db.org/>) website for the sake of exploring the potential

relationships between RRAGB and other genes [19]. We explored the online HPA (<http://www.proteinatlas.org/>) database to validate the RRAGB protein expression of COAD through immunohistochemical staining by antibody HPA003734 [20].

### **Sangerbox website tools**

As a free data analysis platform, the Sangerbox website tools (<http://www.sangerbox.com/tool>) [21,22] was utilized by us to explore the correlations between the RRAGB expression and microsatellite instability (MSI), tumor mutational burden (TMB), tumor neoantigen burden (TNB), tumor microenvironment, tumor immune infiltration, immune checkpoint molecules, immune cells pathway, mismatch repair genes by the Spearman's or the Pearson's method [23–25]. All indicators were visualized by R-packages (<https://www.r-project.org/>) and  $P < 0.001$  was set as the cutoff threshold.

### **Statistical analysis**

All statistical analyses were conducted by using R version 3.5.1 software (<https://www.r-project.org/>). The 'limma' package (<http://www.bioconductor.org/packages/release/bioc/html/limma.html>) was used to compare the RRAGB expression differences between the normal and cancer tissues by Student's *t* test [26]. The Wilcoxon signed-rank test [27] and logistic regression [28] were also performed by us to estimate the associations between the RRAGB expression and clinicopathological variables. Based on the median expression of RRAGB, COAD patients were divided into high-risk groups and low-risk groups. Survival analysis was conducted by Kaplan–Meier (K–M) method with the log-rank test [29]. The receiver operating characteristic (ROC) curves associated with the area under the curve (AUC) values were calculated by the R 'survivalROC' package (<https://cran.r-project.org/web/packages/survivalROC/index.html>). Univariate and multivariate cox hazard regression analyses were applied to identify independent prognostic factors. Furthermore, nomogram was also visualized by the R 'rms' package (<https://cran.r-project.org/web/packages/rms/index.html>) to predict the 1-, 3-, and 5-year survival probabilities and individual predictors. In addition, all *P* values were adopted by a two-

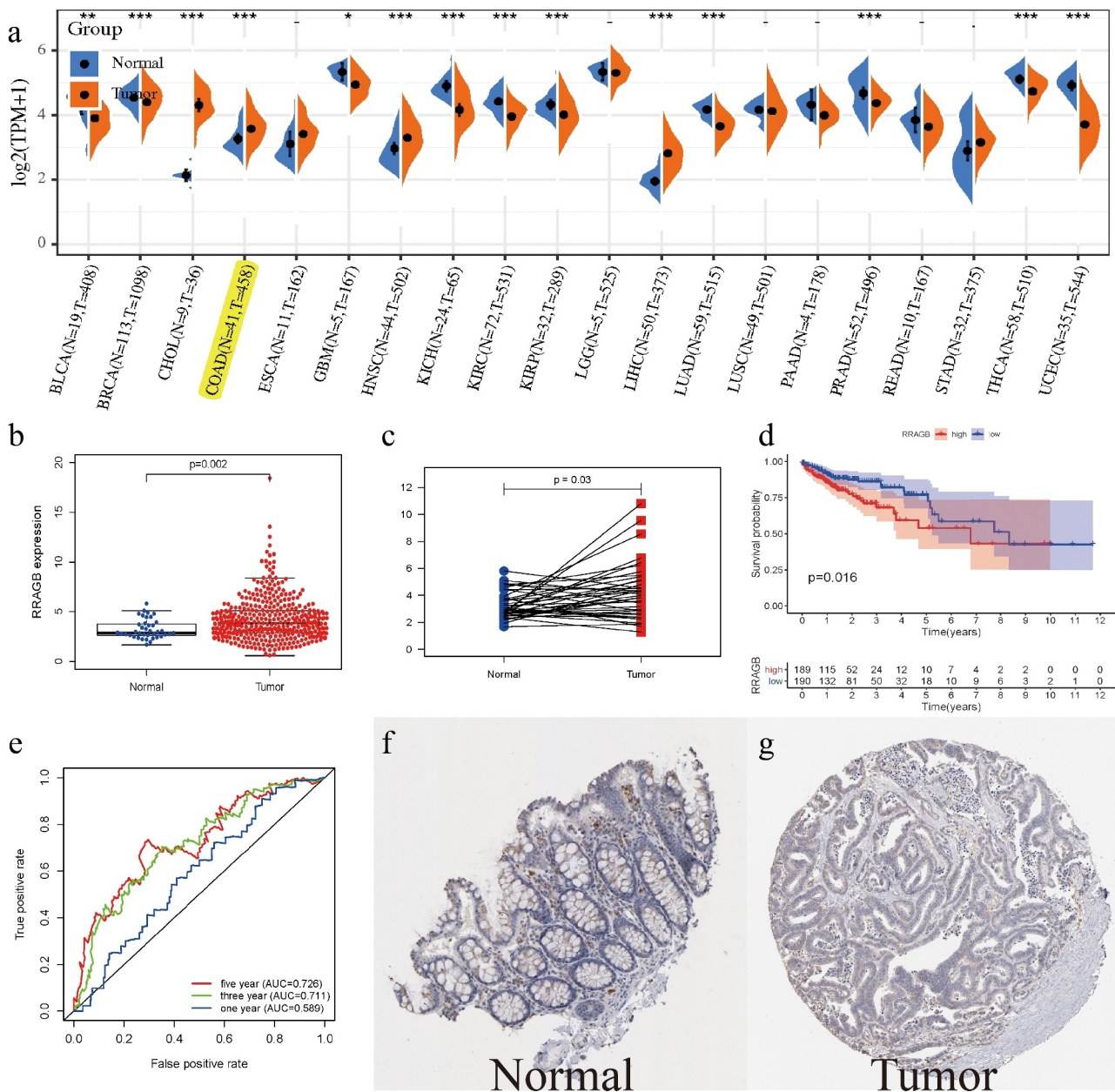
sided test and  $P < 0.05$  was regarded as statistical significance.

## **Results**

By means of TCGA data mining, we aimed to explore the comprehensive roles of RRAGB in COAD and to underline its associations with immunity, hoping to provide a novel candidate gene to improve the prognosis and survival rates of COAD. It was the first time for us to comprehensively explore the expression of RRAGB and its impact on COAD. Our results indicated that RRAGB could serve as an independent prognostic factor for predicting the prognosis of COAD patients and it was significantly associated with MSI, TMB, immunity. GSEA identified five RRAGB-related signaling pathways and RRAGB-based nomogram was also constructed to guide the prognosis of COAD patients. Taken together, RRAGB might be a novel prognostic predictor for COAD.

### **RRAGB expression levels in COAD from TCGA and HPA database**

The RRAGB mRNA expression levels in different cancer types are detailed in Figure 1(a) and we could detect that RRAGB was differently expressed in 14 cancer types including COAD (all  $P < 0.05$ ). We then selected COAD for further analysis. Boxplot ( $N = 39$ ;  $T = 398$ ) and pairwise boxplot ( $N = 39$ ;  $T = 39$ ) confirmed that RRAGB mRNA was up-regulated in COAD tumor tissues, compared with normal tissues (both  $P < 0.05$ ; Figure 1(b,c)). K–M survival analysis indicated that the low-RRAGB group had a much longer OS than the high-RRAGB group, based on its median expression in COAD ( $P = 0.016$ ; Figure 1(d)). Moreover, ROC curves of RRAGB associated with 1-, 3-, and 5-year AUCs were 0.589, 0.711, 0.726, respectively (Figure 1(e)). As displayed in figure 1(f,g), immunohistochemical staining of the HPA database indicated that RRAGB was not detected in normal colon tissues and its expression was medium in COAD tumor tissues. Moreover, we also employed GSE44076 dataset to validate the expression of RRAGB in normal and tumor COAD cells ( $P = 1.92e-07$ ;  $N = 148$ ;  $T = 98$ ; Supplement Figure S1).



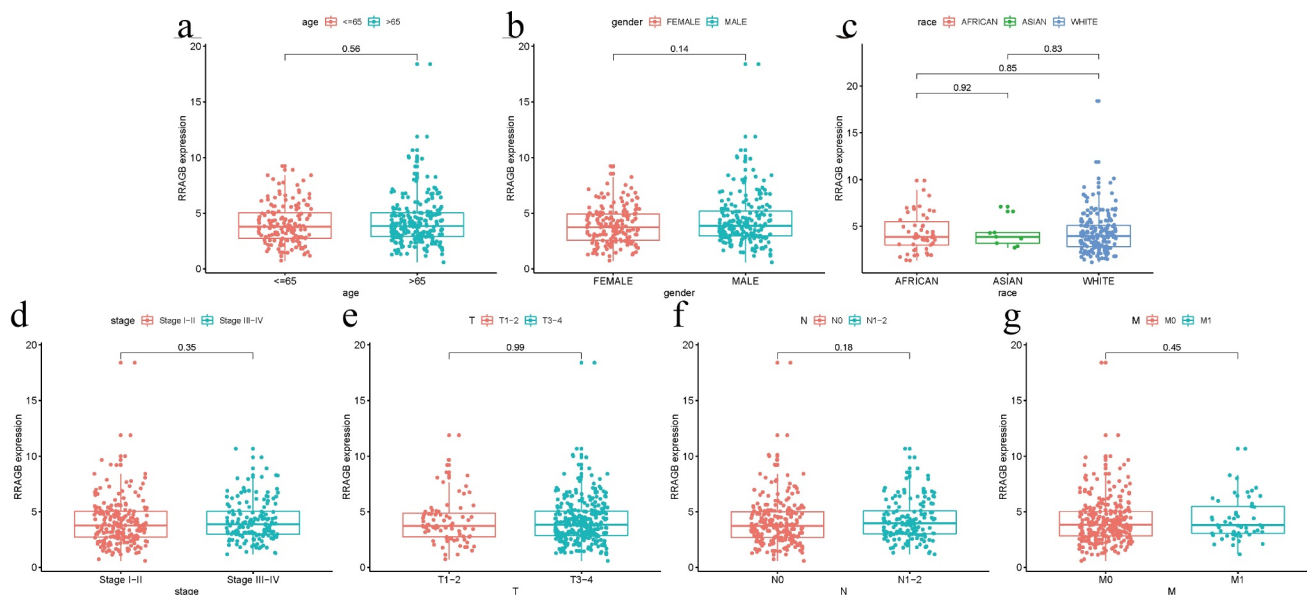
**Figure 1.** The expression levels of RRAGB in COAD tissues; (a) The mRNA expression levels of RRAGB in pan-cancers from the TCGA database; (b) The RRAGB mRNA expression by boxplot in COAD from TCGA (N = 39; T = 398); (c) The RRAGB mRNA expression by pairwise boxplot in COAD from TCGA (N = 39; T = 39); (d) K-M survival analysis of RRAGB in COAD from TCGA; (e) ROC curves and 1-, 3-, and 5-year AUCs of RRAGB in COAD from TCGA; (f-g) The protein expression levels of RRAGB by immunohistochemical staining from the HPA database; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ .

### Associations between RRAGB expression and clinicopathologic variables

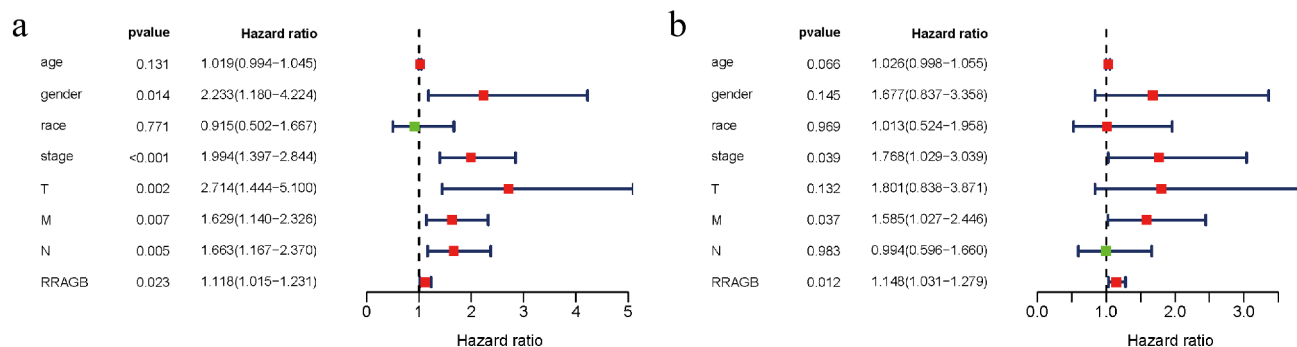
The Wilcoxon signed-rank test and logistic regression were performed by us to estimate the associations between the RRAGB expression and seven clinicopathological variables including age, gender, race, stage, T, N, M. However, no significant associations were observed (all  $P > 0.05$ , Figure 2).

### Univariate and multivariate cox hazard regression analyses

Univariate and multivariate Cox hazard regression analyses were applied by us to identify independent prognostic factors from RRAGB, age, gender, race, stage, T, N and M. Univariate Cox hazard regression analysis showed that gender, stage, T, N, M and RRAGB were all significantly related to OS (all



**Figure 2.** Associations between RRAGB expression and clinicopathologic variables; (a) Associations between RRAGB mRNA expression and age; (b) Associations between RRAGB mRNA expression and gender; (c) Associations between RRAGB mRNA expression and race; (d) Associations between RRAGB mRNA expression and stage; (e) Associations between RRAGB mRNA expression and T stage; (f) Associations between RRAGB mRNA expression and N stage; (g) Associations between RRAGB mRNA expression and M stage.



**Figure 3.** Univariate and multivariate cox hazard regression analyses; (a) Univariate Cox hazard regression analysis of eight clinicopathological variables (RRAGB, age, gender, race, stage, T, N, M) in COAD from TCGA dataset; (b) Multivariate Cox hazard regression analysis of eight clinicopathological variables (RRAGB, age, gender, race, stage, T, N, M) in COAD from TCGA dataset.

$P < 0.05$ ; Figure 3(a) and Table 1). Multivariate Cox hazard regression analysis presented that stage, M and RRAGB were markedly linked to the OS (all  $P < 0.05$ ; Figure 3(b) and Table 1). In other words, stage, M and RRAGB were all independent prognostic factors for COAD.

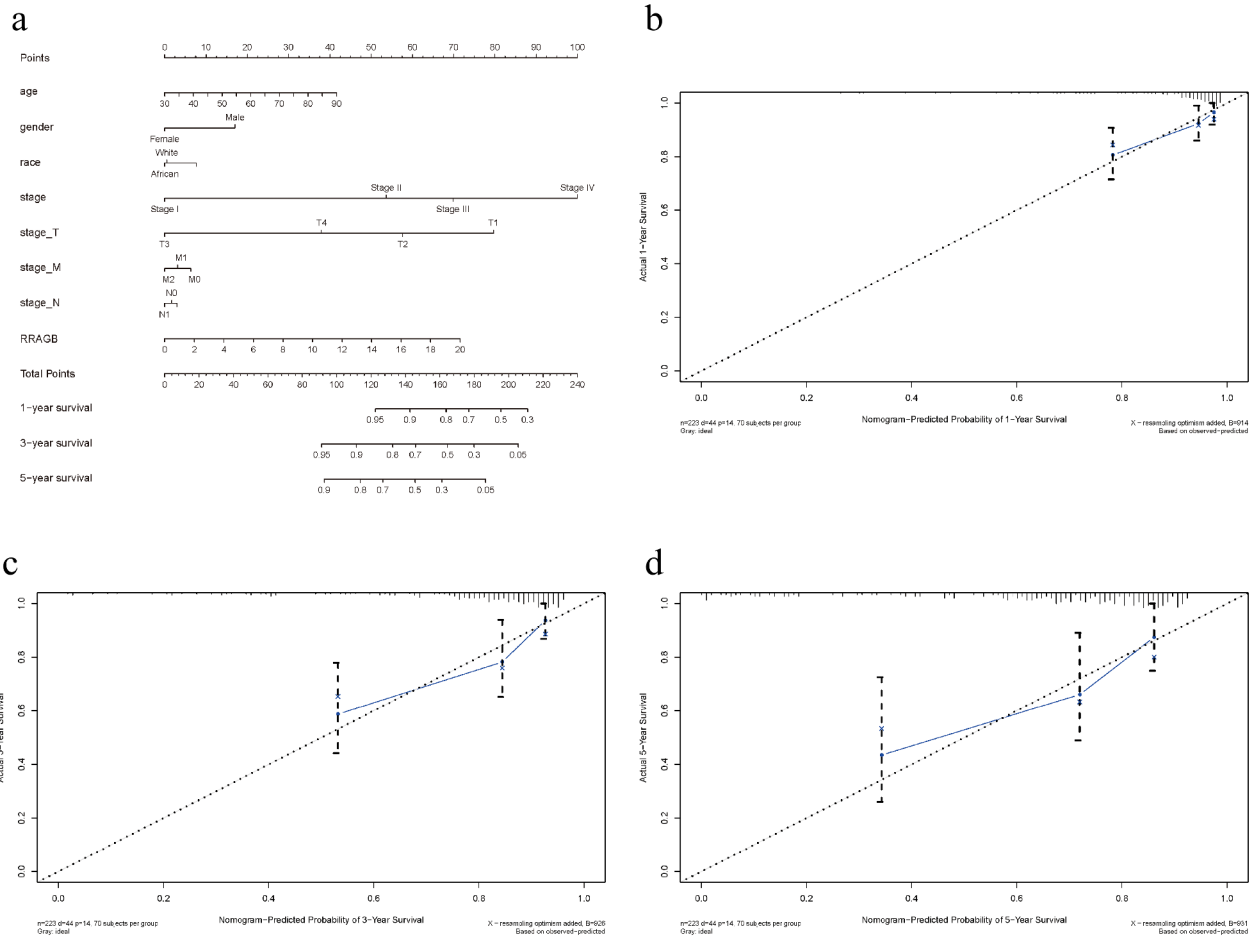
### Construction of RRAGB based nomogram

A nomogram was constructed by the R 'rms' package to intuitively visualize the associations between

eight clinicopathological variables (RRAGB, age, gender, race, stage, T, N and M) and 1-, 3-, 5-year survival probabilities of OS (Figure 4(a)). As summarized in Table 2, C-index and 1-, 3-, 5-year AUCs of RRAGB-based nomogram were 0.872, 0.748, 0.799 and 0.791, showing a moderate prediction accuracy. As displayed in Figure 4(b-d), 1-, 3-, and 5-year calibration curves indicated the consistency of our results and the predictive values, indicating satisfactory performance for this RRAGB-based nomogram.

**Table 1.** Univariate and multivariate cox hazard regression analyses of COAD in TCGA.

Characteristics	Univariate analysis				Multivariate analysis			
	HR	HR.95 L	HR.95 H	pvalue	HR	HR.95 L	HR.95 H	pvalue
<b>Age</b>	1.019482	0.994269	1.045334	0.131014	1.026415	0.998231	1.055396	0.06646
<b>Gender</b>	2.23311	1.180471	4.224399	<b>0.01351</b>	1.676718	0.837335	3.357536	0.144605
<b>Race</b>	0.914905	0.502185	1.666818	0.771367	1.013196	0.524169	1.958468	0.9689
<b>Stage</b>	1.993568	1.397495	2.843882	<b>0.00014</b>	1.768459	1.029165	3.038821	<b>0.03901</b>
<b>T</b>	2.713683	1.444052	5.099591	<b>0.00192</b>	1.800819	0.837723	3.871144	0.131941
<b>M</b>	1.628803	1.14046	2.326253	<b>0.0073</b>	1.585269	1.027464	2.445903	<b>0.0373</b>
<b>N</b>	1.663132	1.167274	2.369631	<b>0.00486</b>	0.994441	0.59583	1.659724	0.982982
<b>RRAGB</b>	1.118193	1.015388	1.231407	<b>0.02319</b>	1.148259	1.030556	1.279405	<b>0.01223</b>



**Figure 4.** RRAGB based nomogram and evaluation; (a) Nomogram based on eight clinicopathological variables (RRAGB, age, gender, race, stage, T, N, M) in COAD from TCGA dataset; (b) 1-year calibration curve; (c) 3-year calibration curve; (d) 5-year calibration curve.

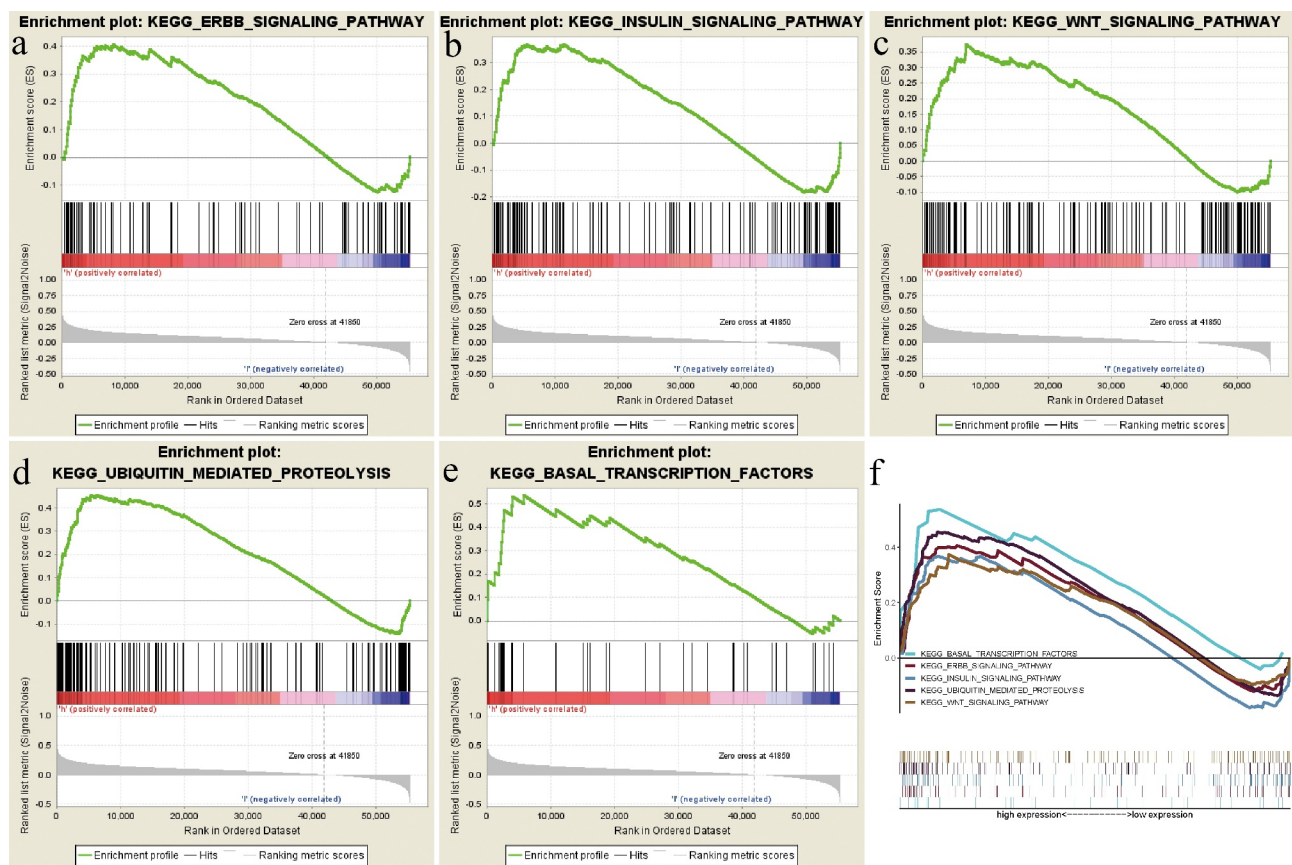
**Table 2.** C-index and 1-, 3-, 5-year AUCs of RRAGB-based nomogram.

	1-year	3-year	5-year	C-index
<b>AUC</b>	0.748	0.799	0.791	0.872

**RRAGB-related signaling pathways according to GSEA**

GSEA was by us performed to reveal significant survival differences between the high-RRAGB and

low-RRAGB expression groups to discover significant critical biological pathways, based on the the ‘c2.cp.kegg.v6.2.symbols.gmt’ gene set from the MSigDB. Our results indicated that the high-RRAGB expression phenotype was significantly associated with Basal transcription factors, Ubiquitin mediated proteolysis, Insulin, Wnt, Erbb signaling pathways, with the threshold of



**Figure 5.** Gene set enrichment analysis (GSEA) results based on the RRAGB mRNA expression in COAD from TCGA dataset; (a) Significantly enriched pathway of ErbB signaling pathway; (b) Significantly enriched pathway of Insulin signaling pathway; (c) Significantly enriched pathway of Wnt signaling pathway; (d) Significantly enriched pathway of Ubiquitin mediated proteolysis; (e) Significantly enriched pathway of Basal transcription factors; (f) All of these five significantly enriched signaling pathways.

**Table 3.** Gene set enrichment analysis results.

MSigDB collection	Gene set name	NES	NOM p-val	FDR q-val
c2.cp.kegg.v7.1.symbols.gmt	KEGG_ERBB_SIGNALING_PATHWAY	1.524	0.040	0.603
	KEGG_INSULIN_SIGNALING_PATHWAY	1.516	0.029	0.501
	KEGG_BASAL_TRANSCRIPTION_FACTORS	1.675	0.029	1.000
	KEGG_UBIQUITIN_MEDIATED_PROTEOLYSIS	1.759	0.012	1.000
	KEGG_WNT_SIGNALING_PATHWAY	1.535	0.037	0.764

NES >1.5 and normal p values <0.05 (Figure 5 and Table 3).

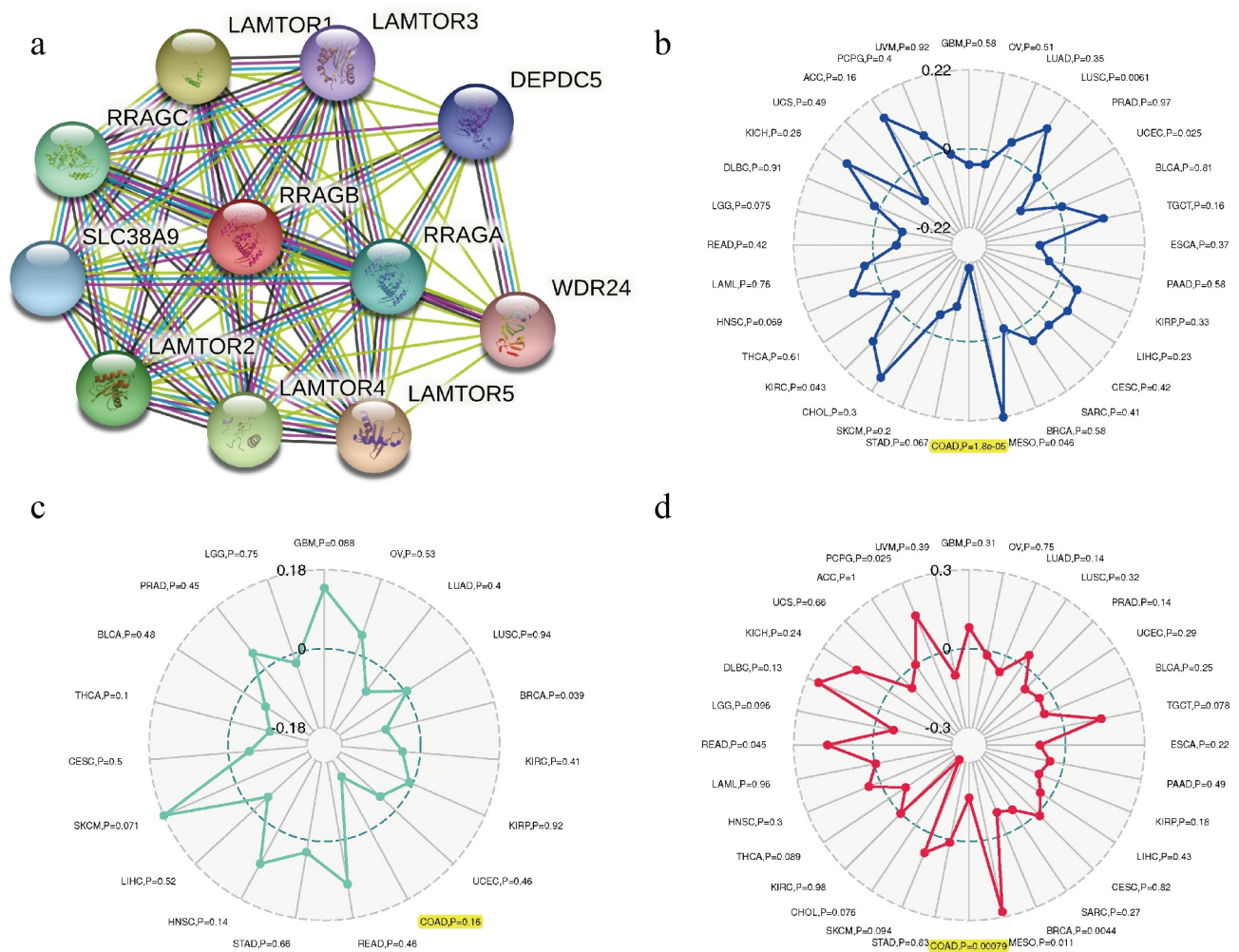
### Associations between RRAGB and PPI network, MSI, TNB, TMB in COAD

A PPI network analysis was carried out with the help of the online STRING (<https://string-db.org/>) website to explore the potential relationships between RRAGB and other genes in COAD (Figure 6(a)). By means of the Pearson's method, we calculated the correlations between the RRAGB expression and MSI, TNB, TMB, with the help of the Sangerbox website tools (<http://www.sangerbox.com/tool>).

Radar maps indicated that the RRAGB expression was dramatically linked to MSI ( $P = 1.8 \times 10^{-5}$ ) and TMB ( $P = 0.00079$ ) in COAD, whereas it was not associated with TNB ( $P = 0.16$ ).

### Associations between RRAGB and tumor immune infiltration, SCNA, tumor microenvironment in COAD

With the help of the Sangerbox website tools, we calculated the correlations between the RRAGB expression and tumor immune infiltration by means of the Spearman's method. The RRAGB mRNA expression was markedly related to B cells,



**Figure 6.** Associations between RRAGB and PPI network, MSI, TNB, TMB in COAD; (a) PPI network based on RRAGB expression; (b) Associations between RRAGB expression and MSI in COAD; (c) Associations between RRAGB expression and TNB in COAD; (d) Associations between RRAGB expression and TMB in COAD.

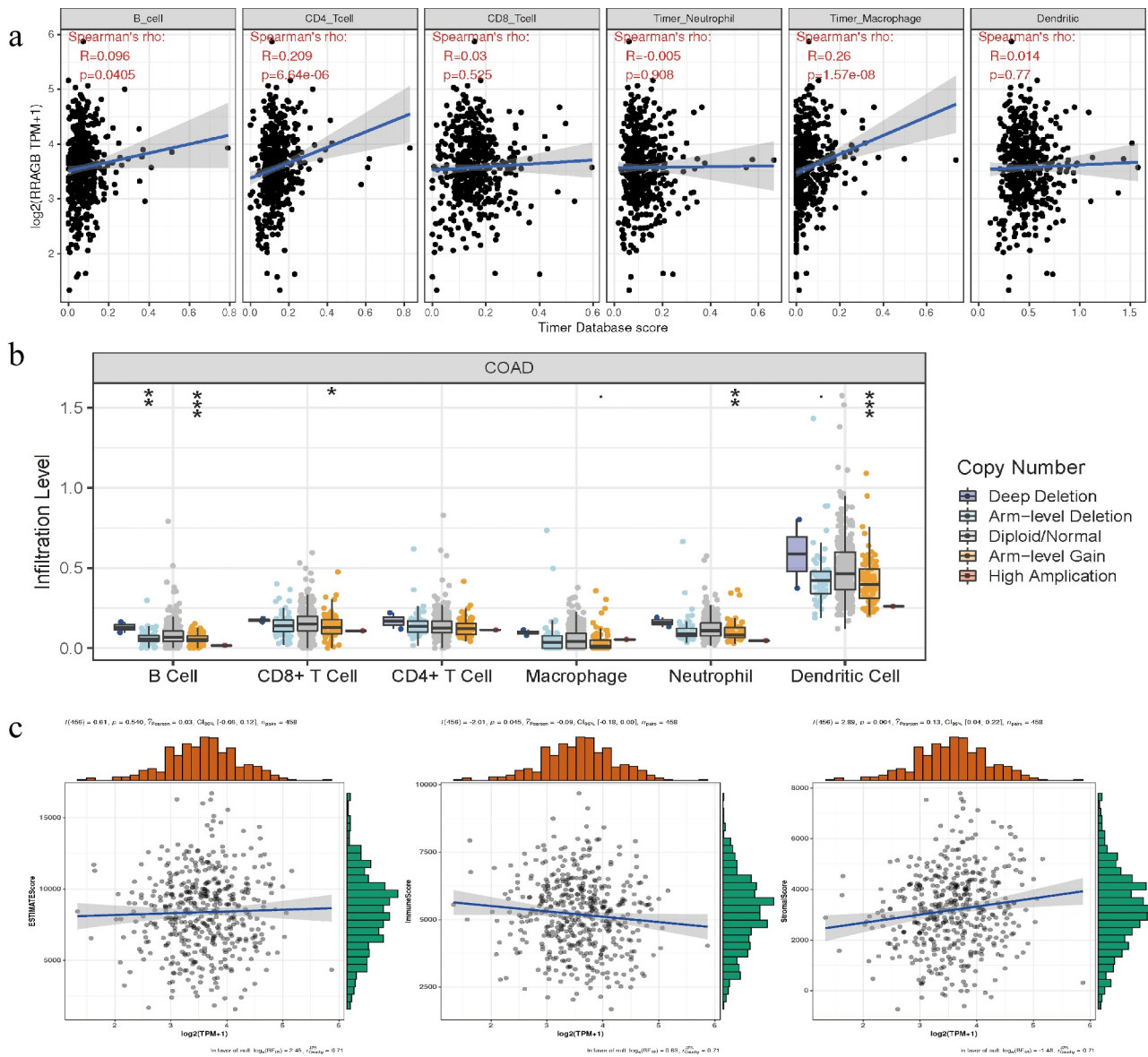
CD4 + T cells and Macrophage cells infiltration (all  $P < 0.05$ , Figure 7(a)). SCNA module provided the correlations between tumor immune infiltration levels among COAD and different somatic copy number alterations for RRAGB by the Wilcoxon rank-sum test (Figure 7(b)). With the Pearson's method, the RRAGB mRNA expression was significantly associated with ImmuneScore and StromalScore (both  $P < 0.05$ ), whereas it was not linked to ESTIMATEScore ( $P = 0.540$ , Figure 7(c)).

#### **Correlations between RRAGB and immune checkpoint molecules, immune cells pathway, mismatch repair genes in COAD**

By means of the Pearson's method, we calculated the correlations between the RRAGB expression

and immune checkpoint molecules, immune cells pathway, mismatch repair genes, with the help of the Sangerbox website tools. As for immune checkpoint molecules, the RRAGB mRNA expression was significantly associated with TNFSF4, TNFSF9, TNFSF18, TMIGD2, TNFRSF14, TNFRSF18 in COAD (all  $P < 0.05$ , Figure 8(a)). In terms of immune cells pathways, the RRAGB mRNA expression was markedly related to Type 17 T helper cell, Type 2 T helper cell, Neutrophil, Monocyte, Memory B cell, CD56dim natural killer cell, Activated dendritic cell and so on in COAD (all  $P < 0.05$ , Figure 8(b)). Co-expression analysis of RRAGB and mismatch repair genes indicated that RRAGB was dramatically linked to MLH1, MSH2, MSH6, PMS2, EPCAM in COAD (all  $P < 0.001$ , Figure 8(c)).



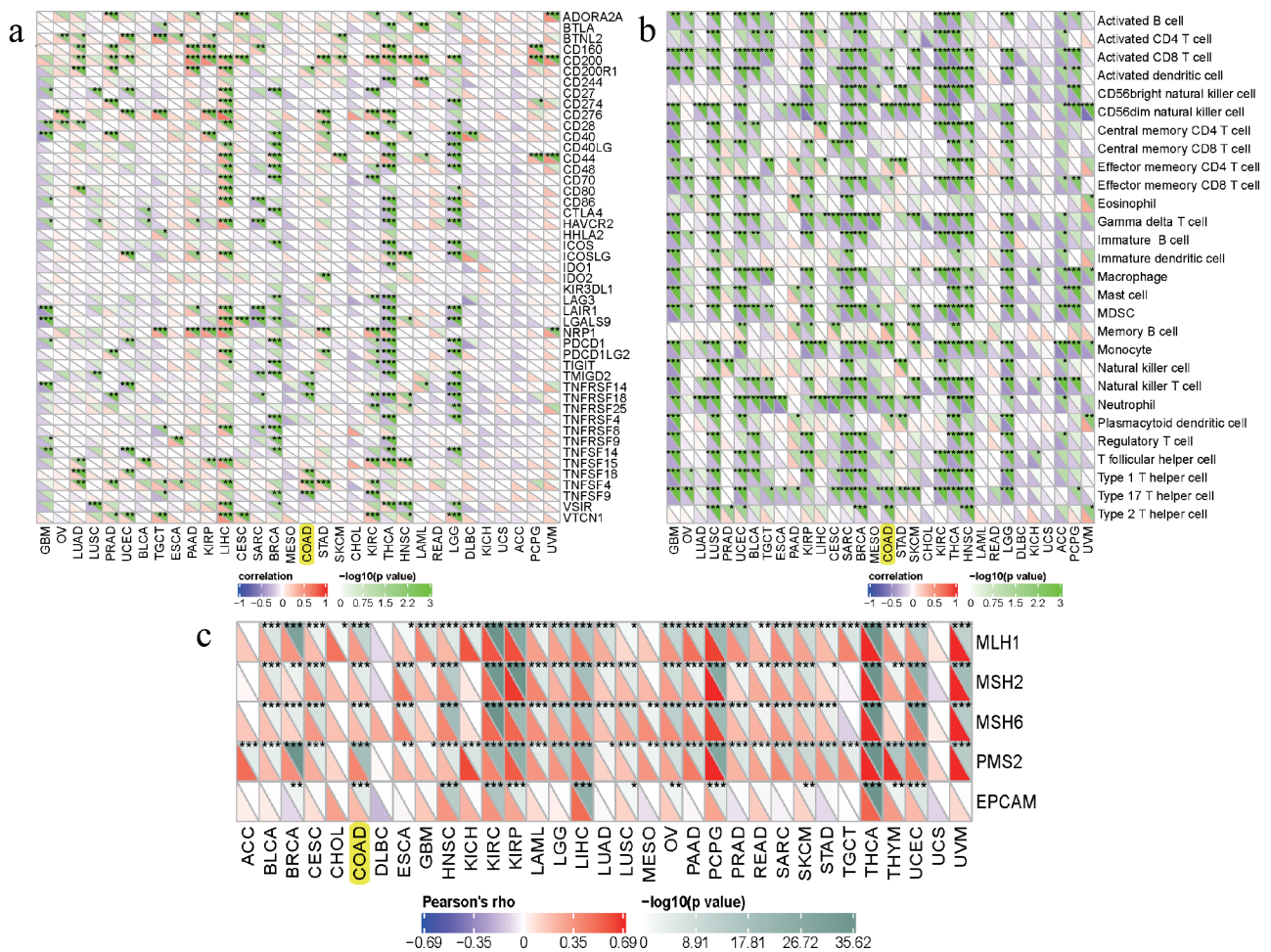


**Figure 7.** Associations between RRAGB expression and tumor immune infiltration, SCNA, tumor microenvironment in COAD; (a) Associations between RRAGB expression and tumor immune infiltration in COAD from TCGA dataset; (b) Associations between RRAGB expression and SCNA in COAD from TCGA dataset; (c) Associations between RRAGB expression and tumor microenvironment in COAD from TCGA dataset; \* $P < 0.05$ ; \*\* $P < 0.05$ ; \*\*\* $P < 0.001$ .

## Discussion

Currently, no studies have reported the definite roles of RRAGB in human cancers and it was the first time for us to comprehensively explore the expression of RRAGB and its impact on COAD. In this article, our results indicated that RRAGB was not only differently expressed in COAD, but also significantly associated with OS. Further univariate and multivariate Cox hazard regression analyses indicated that stage, M and RRAGB could all serve as independent prognostic factors for

COAD. As reported by previous researches, Bunte et al. found that RRAGB was more prevalent in periodontitis patients than healthy controls and might be associated with periodontitis clinical manifestations [30]. Shi et al. successfully constructed a signature based on CHMP4C, FOXO1, RRAGB, and effectively predicted the cervical cancer patients' prognosis [13]. Xie et al. also constructed a six-gene model including RRAGB and it could function as an independent prognostic factor to predict the non-small-cell lung cancer patients' OS [14].



**Figure 8.** Correlations between RRAGB expression and immune checkpoint molecules, immune cells pathway, mismatch repair genes in COAD; (a) Correlations between RRAGB expression and immune checkpoint molecules in COAD from TCGA dataset; (b) Correlations between RRAGB expression and immune cells pathway in COAD from TCGA dataset; (c) Correlations between RRAGB expression and mismatch repair genes in COAD from TCGA dataset; \* $P < 0.05$ ; \*\* $P < 0.05$ ; \*\*\* $P < 0.001$ .

In this article, our results indicated that the high-RRAGB expression phenotype was significantly associated with Basal transcription factors, Ubiquitin-mediated proteolysis, Insulin, Wnt, Erbb signaling pathways. As a useful tool, GSEA had been applied by various studies to reveal significant survival differences between the high- and low-gene expression groups to discover significant critical biological pathways [31]. Shuai et al found that eight signaling pathways (apoptosis, cell cycle, ErbB, MAPK, mTOR, Notch, p53 and TGF- $\beta$  pathways) were most significantly enriched in the high-CDCA8 phenotype, according to the GSEA [11]. Liu et al. reveal that ZNF132 could participate in multiple biological pathways containing the regulation of glycolysis and cell cycle, based on the results of GSEA [32]. Consistent with our results, all of these three signaling pathways had

been reported to be markedly related to colon cancer [33–35]. As reported by previous articles, Wnt and ERBB signaling pathway played important roles in cancer [36,37]. Yan et al. reported that LINC00261 could regulate miR-324-3p and the Wnt signaling pathway to repress colon cancer progression [38]. Wu et al. showed that tankyrase 1 inhibitor (XAV939) could increase chemosensitivity in colon cancer cells through inhibiting the Wnt signaling pathway [39]. As for Erbb signaling pathway, He et al. reported that YAP could form an autocrine loop with the ERBB pathway to regulate the initiation and progression of ovarian cancer [40]. Grimont revealed that SOX9 could regulate the ERBB signaling pathway in the development of pancreatic cancer [41].

Due to the ability of reducing statistically predictive models into single numerical estimated

probabilities, nomograms had been widely used for cancer prognosis and clinical decision-making [42]. Wang et al found that nomogram based on serum cystatin C was beneficial to evaluate acute kidney injury (AKI) possibilities and avoid its occurrences [43]. Liu et al. successfully establish a TP53-associated nomogram and exhibited excellent efficacy in predicting the OS prognosis of pancreatic cancer patients [44]. In this article, we also constructed a nomogram to intuitively visualize the associations between eight clinicopathological variables (RRAGB, age, gender, race, stage, T, N and M) and 1-, 3-, 5-year survival probabilities of OS. After evaluated by C-index, ROC and calibration curves, our established RRAGB-based nomogram displayed a satisfactory performance.

As reported by previous studies, TMB, TNB, MSI could serve as biomarkers related to the immune checkpoint inhibitors' efficacy and survival prognosis [45,46]. Therefore, we calculated the correlations between the RRAGB expression and MSI, TNB, TMB by means of the pearson's method and results indicated that the RRAGB expression was dramatically linked to MSI and TMB in COAD, whereas it was not associated with TNB. The dynamic characteristics of the tumor immune infiltration, tumor microenvironment, immune checkpoint molecules and immune cells pathway were vital for immunotherapy and played essential roles in tumorigenesis and progression [47–49]. In our article, we calculated the correlations between the RRAGB expression and tumor immune infiltration by the spearman's method and found that the RRAGB mRNA expression was markedly related to B cells, CD4 + T cells and Macrophage cells infiltration. Through the pearson's method, the RRAGB mRNA expression was significantly associated with ImmuneScore and StromalScore, whereas it was not linked to ESTIMATEScore. In terms of immune checkpoint molecules and immune cells pathway, the RRAGB mRNA expression was significantly associated with TNFSF4, TNFSF9, TNFSF18, TMIGD2, TNFRSF14, TNFRSF18, Type 17T helper cell, Type 2T helper cell, neutrophil, Monocyte, Memory B cell, CD56dim natural killer cell, activated dendritic cell and so on in COAD by means of the pearson's method. All of these indicated the strong associations between RRAGB and immunity.

The strength of this article was that it was the first time for us to comprehensively explore the expression of RRAGB and its impact on COAD. Univariate and multivariate cox hazard regression analyses indicated that RRAGB could serve as independent prognostic factors for COAD and GSEA identified RRAGB-related signaling pathways. RRAGB was revealed to be significantly associated with MSI, TMB, immunity and we also constructed an RRAGB-based nomogram, having a satisfactory performance. There were several limitations too. Firstly, clinical information was merely collected from TCGA dataset, which was limited and insufficient. Treatment information in TCGA was absent. Secondly, we mainly paid attention to the bioinformatics analysis of RRAGB, without experimental validation. We would further carry out in vivo and in vitro experiments to verify our results.

## Conclusions

Overall, our results revealed that RRAGB could be a prognostic biomarker for COAD in terms of OS. Univariate and multivariate Cox hazard regression analyses indicated that RRAGB could serve as independent prognostic factors for COAD and GSEA identified RRAGB-related signaling pathways. Moreover, RRAGB was found to be significantly associated with MSI, TMB, and immunity. Last but not least, we also constructed an RRAGB-based nomogram, having a satisfactory performance. Further researches with more sample sizes and experiments in vivo and in vitro were required to verify our findings.

## Highlights

- It was the first time for us to comprehensively explore the expression of RRAGB and its impact on COAD.
- RRAGB could serve as an independent prognostic factor for predicting the prognosis of COAD patients.
- RRAGB was significantly associated with MSI, TMB, immunity.
- GSEA identified five RRAGB-related signaling pathways.

- RRAGB-based nomogram was also constructed to guide the prognosis of COAD patients.

## Acknowledgements

We would like to thank the researchers and study participants for their contributions.

## Availability of data and material

All data and clinical information could be obtained from The Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>).

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Author's contribution

SR.Z: Protocol/project development; Y.Y, WJ.W: Data collection or management; J.Z, JY.S: Data analysis and manuscript revision; JJ.X, QQ.L: Manuscript writing/editing. All authors have read and approved the final manuscript.

## References

- [1] Siegel RL, Miller KD, Fuchs HE, et al. Cancer statistics, 2021. *CA Cancer J Clin.* 2021;71(1):7–33.
- [2] Chang K, Yuan C, Liu X. A new RBPs-related signature predicts the prognosis of colon adenocarcinoma patients. *Front Oncol.* 2021;11:627504.
- [3] Punt CJ, Koopman M, Vermeulen L. From tumour heterogeneity to advances in precision treatment of colorectal cancer. *Nat Rev Clin Oncol.* 2017;14(4):235–246.
- [4] Brown KGM, Solomon MJ. Progress and future direction in the management of advanced colorectal cancer. *Br J Surg.* 2018;105(6):615–617.
- [5] Bresalier RS, Kopetz S, Brenner DE. Blood-based tests for colorectal cancer screening: do they threaten the survival of the FIT test? *Dig Dis Sci.* 2015;60(3):664–671.
- [6] Siegel R, DeSantis C, Virgo K, et al. Cancer treatment and survivorship statistics, 2012. *CA Cancer J Clin.* 2012;62(4):220–241.
- [7] Liang Y, Su Q, Wu X. Identification and validation of a novel six-gene prognostic signature of stem cell characteristic in colon cancer. *Front Oncol.* 2020;10:571655.
- [8] Schürmann A, Brauers A, Maßmann S, et al. Cloning of a novel family of mammalian GTP-binding proteins (RagA, RagBs, RagB1) with remote similarity to the Ras-related GTPases. *J Biol Chem.* 1995;270(48):28982–28988.
- [9] Sancak Y, Peterson TR, Shaul YD, et al. The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. *Science.* 2008;320(5882):1496–1501.
- [10] Liu S, Wang Y, Miao C, et al. High expression of CDCA7 predicts poor prognosis for clear cell renal cell carcinoma and explores its associations with immunity. *Cancer Cell Int.* 2021;21(1):140.
- [11] Shuai Y, Fan E, Zhong Q, et al. CDCA8 as an independent predictor for a poor prognosis in liver cancer. *Cancer Cell Int.* 2021;21(1):159.
- [12] Chen Q, Qiu B, Zeng X, et al. Identification of a tumor microenvironment-related gene signature to improve the prediction of cervical cancer prognosis. *Cancer Cell Int.* 2021;21(1):182.
- [13] Shi H, Zhong F, Yi X, et al. Application of an autophagy-related gene prognostic risk model based on TCGA database in cervical cancer. *Front Genet.* 2020;11:616998.
- [14] Xie H, Xie C. A six-gene signature predicts survival of adenocarcinoma type of non-small-cell lung cancer patients: a comprehensive study based on integrated analysis and weighted gene coexpression network. *Biomed Res Int.* 2019;2019:4250613.
- [15] Liu J, Xu F, Cheng W, et al. Identification and verification of a ten-gene signature predicting overall survival for ovarian cancer. *Exp Cell Res.* 2020;395(2):112235.
- [16] Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47.
- [17] Liberzon A, Birger C, Thorvaldsdóttir H, et al. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst.* 2015;1(6):417–425.
- [18] Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102(43):15545–15550.
- [19] Szklarczyk D, Franceschini A, Kuhn M, et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res.* 2011;39:D561–8.
- [20] Uhlén M, Fagerberg L, Hallström BM, et al. Proteomics. Tissue-based map of the human proteome. *Science.* 2015;347(6220):1260419.
- [21] Hu J, Qiu D, Yu A, et al. YTHDF1 Is a potential pan-cancer biomarker for prognosis and immunotherapy. *Front Oncol.* 2021;11:607224.
- [22] Zhang L, Liu Z, Dong Y, et al. E2F2 drives glioma progression via PI3K/AKT in a PFKFB4-dependent manner. *Life Sci.* 2021;276:119412.
- [23] Hundal J, Carreno BM, Petti AA, et al. pVAC-Seq: a genome-guided in silico approach to identifying tumor neoantigens. *Genome Med.* 2016;8(1):11.

- [24] Yoshihara K, Shahmoradgoli M, Martínez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun.* **2013**;4(1):2612.
- [25] Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* **2015**;12(5):453–457.
- [26] Gauvreau K, Pagano M. Student's t test. *Nutrition.* **1993**;9(4):386.
- [27] Crichton N. Information point: wilcoxon signed rank test. *J Clin Nurs.* **2000**;9(4):584.
- [28] Schober P, Vetter TR. Logistic regression in medical research. *Anesth Analg.* **2021**;132(2):365–366.
- [29] Schober P, Vetter TR. Kaplan-meier curves, log-rank tests, and cox regression for time-to-event data. *Anesth Analg.* **2021**;132(4):969–970.
- [30] Bunte K, Kuhn C, Walther C, et al. Clinical significance of ragA, ragB, and PG0982 genes in *Porphyromonas gingivalis* isolates from periodontitis patients. *Eur J Oral Sci.* **2021**; 129(3):e12776.
- [31] Ma Y, Sun S, Shang X, et al. Integrative differential expression and gene set enrichment analysis using summary statistics for scRNA-seq studies. *Nat Commun.* **2020**;11(1):1585.
- [32] Liu Z, Liu J, Liu R, et al. Downregulated ZNF132 predicts unfavorable outcomes in breast Cancer via Hypermethylation modification. *BMC Cancer.* **2021**;21(1):367.
- [33] Szmida E, Karpiński P, Leszczynski P, et al. Aberrant methylation of ERBB pathway genes in sporadic colorectal cancer. *J Appl Genet.* **2015**;56(2):185–192.
- [34] Liu S, Shi Y, Liu Z, et al. Circular RNA SMARCA5 functions as an anti-tumor candidate in colon cancer by sponging microRNA-552. *Cell Cycle.* **2021**:1–13. DOI:10.1080/15384101.2021.1949840.
- [35] Keku TO, Vidal A, Oliver S, et al. Genetic variants in IGF-I, IGF-II, IGFBP-3, and adiponectin genes and colon cancer risk in African Americans and Whites. *Cancer Causes Control.* **2012**;23(7):1127–1138.
- [36] Feng L, Huang S, An G, et al. Identification of new cancer stem cell markers and signaling pathways in HER-2-positive breast cancer by transcriptome sequencing. *Int J Oncol.* **2019**;55(5):1003–1018.
- [37] Wang Y, Cong R, Liu S, et al. Decreased expression of METTL14 predicts poor prognosis and construction of a prognostic signature for clear cell renal cell carcinoma. *Cancer Cell Int.* **2021**;21(1):46.
- [38] Yan D, Liu W, Liu Y, et al. LINC00261 suppresses human colon cancer progression via sponging miR-324-3p and inactivating the Wnt/ $\beta$ -catenin pathway. *J Cell Physiol.* **2019**;234(12):22648–22656.
- [39] Wu X, LUO F, Li J, et al. Tankyrase 1 inhibitor XAV939 increases chemosensitivity in colon cancer cell lines via inhibition of the Wnt signaling pathway. *Int J Oncol.* **2016**;48(4):1333–1340.
- [40] He C, Lv X, Hua G, et al. YAP forms autocrine loops with the ERBB pathway to regulate ovarian cancer initiation and progression. *Oncogene.* **2015**;34(50):6040–6054.
- [41] Grimont A, Pinho AV, Cowley MJ, et al. SOX9 regulates ERBB signalling in pancreatic cancer development. *Gut.* **2015**;64(11):1790–1799.
- [42] Iasonos A, Schrag D, Raj GV, et al. How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol.* **2008**;26(8):1364–1370.
- [43] Wang RR, He M, Gui X, et al. A nomogram based on serum cystatin C for predicting acute kidney injury in patients with traumatic brain injury. *Ren Fail.* **2021**;43(1):206–215.
- [44] Liu X, Chen B, Chen J, et al. A novel tp53-associated nomogram to predict the overall survival in patients with pancreatic cancer. *BMC Cancer.* **2021**;21(1):335.
- [45] Wang B, Zhong J-L, Li H-Z, et al. Diagnostic and therapeutic values of PMEPA1 and its correlation with tumor immunity in pan-cancer. *Life Sci.* **2021**;277:119452.
- [46] Zeng D, Ye Z, Wu J, et al. Macrophage correlates with immunophenotype and predicts anti-PD-L1 response of urothelial cancer. *Theranostics.* **2020**;10(15):7002–7014.
- [47] Jiao S, Subudhi SK, Aparicio A, et al. Differences in tumor microenvironment dictate T helper lineage polarization and response to immune checkpoint therapy. *Cell.* **2019**;179(5):1177–1190.e13.
- [48] Neal JT, Li X, Zhu J, et al. Organoid modeling of the tumor immune microenvironment. *Cell.* **2018**;175(7):1972–1988.e16.
- [49] Kaderbhai C, Tharin Z, Ghiringhelli F. The role of molecular profiling to predict the response to immune checkpoint inhibitors in lung cancer. *Cancers (Basel).* **2019**;11(2):201.