Translational Oncology 28 (2023) 101609

Contents lists available at ScienceDirect

Translational Oncology

journal homepage: www.elsevier.com/locate/tranon

Original Research

Development of novel DNAJB6-KIAA1522-p-mTOR three-protein prognostic prediction models for CRC

Yu-Juan Jiang [a,c,1], Tong-Tong Zhang [a,b,1], Yi-Qing Zhu [a], Hong-Qing Cai [a], Chen Chang [a], Jia-Jie Hao [a], Yan Cai [a], Ming-Rong Wang [a], Jian-Wei Liang [c,*], Yu Zhang [a,*]

[a] State Key Laboratory of Molecular Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100021, China
[b] Medical Research Center, The Third People's Hospital of Chengdu, Chengdu, 610031, China
[c] Department of Colorectal Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100021, China

ARTICLE INFO

Keywords:
Colorectal cancer
Prognostic biomarkers
Molecular classification
KIAA1522
DNAJB6
p-mTOR

ABSTRACT

Background: To evaluate the prognostic value of DNAJB6, KIAA1522, and p-mTOR expression for colorectal cancer (CRC) and to develop effective prognostic models for CRC patients.
Methods: The expression of DNAJB6, KIAA1522, and p-mTOR (Ser2448) was detected using immunohistochemistry in 329 CRC specimens. The prognostic values of the three proteins in the training cohort were assessed using Kaplan-Meier curves and univariate and multivariate Cox proportional hazards models. Prediction nomogram models integrating the three proteins and TNM stage were constructed. Subsequently, calibration curves, receiver operating characteristic (ROC) curves, the concordance index (C-index), and decision curve analysis (DCA) were used to evaluate the performance of the nomograms in the training and validation cohorts.
Results: The three proteins DNAJB6, KIAA1522, and p-mTOR were significantly overexpressed in CRC tissues (each $P < 0.01$), and their expression was an independent prognostic factor for overall survival (OS) and disease-free survival (DFS) (each $P < 0.05$). The area under the ROC curves (AUC) and C-index values were approximately 0.7. Additionally, the calibration curves showed that the predicted values and the actual values fit well. Furthermore, DCA curves indicated that the clinical value of the nomogram models was higher than that of TNM stage. Overall, the novel prediction models have good discriminability, sensitivity, specificity and clinical utility.
Conclusion: The nomograms containing DNAJB6, KIAA1522, and p-mTOR may be promising models for predicting postoperative survival in CRC.

Introduction

Colorectal cancer (CRC) was the third most prevalent malignant tumor and the second leading cause of malignant disease death worldwide in 2020 [1]. In China, it is estimated that there were 408000 new CRC cases and 195600 deaths from CRC in 2016, ranking second in morbidity and fourth in mortality [2]. Of note, CRC incidence has been rising rapidly in China and has become a significant cause of cancer death [2].

Although recent advances in surgery, chemotherapy, radiotherapy, targeted therapy, and immunotherapy have improved the prognosis of CRC patients, the 5-year survival rate is only approximately 55% [3]. Accurate prognosis prediction is essential for effective and individualized treatment. At present, the outcomes of CRC patients are routinely determined using the tumor-node-metastasis (TNM) staging system. However, recurrence and mortality might vary greatly even among CRC patients with similar clinicopathological characteristics because of the high molecular heterogeneity in CRC [4,5]. Thus, it is necessary to develop simple, user-friendly, and reliable molecular markers to classify the prognosis of CRC patients [6].

Various biomarkers or classifiers, such as DNA, RNA (including circulating tumor DNA and RNA), protein, epigenetic (such as DNA

* Corresponding author at: National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, 17 Panjiayuan Nanli, Chaoyang District, Beijing 100021, China.
E-mail addresses: liangjw1976@163.com (J.-W. Liang), zhangyu909@126.com (Y. Zhang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.tranon.2022.101609
Received 11 September 2022; Received in revised form 28 November 2022; Accepted 19 December 2022

methylation) and immunologic biomarkers, have recently been developed to predict the prognosis of patients with CRC, particularly their recurrence risk [7–13]. However, additional prospective clinical trials are needed to evaluate the predictive accuracy and efficiency of these putative biomarkers. Compared to sequencing-based methods, immunohistochemistry (IHC) is simpler, easier to perform, and less expensive, and it is widely used in biomarker research [14]. In our previous study, we discovered that several proteins were markedly overexpressed in CRC tissues but scarcely detected in the adjacent normal colorectal epithelium. Among them, the aberrant expression of several proteins was significantly correlated with poor CRC patient prognosis. Based on these findings, we selected three proteins, namely, p-mTOR (Ser 2448), DNAJB6, and KIAA152, which were closely related to metastatic potential, to construct prognostic prediction models.

Mammalian/mechanistic target of rapamycin (mTOR) is a serine/threonine-protein kinase belonging to the PI3K-related protein kinase (PIKK) family [15]. Abnormal activation of the PI3K/AKT/mTOR pathway has been reported in various cancer types [16–19]. Its hyperactivation has been closely associated with the prognosis of several cancers, including CRC [20]. DnaJ/Hsp40 homolog, subfamily B, member 6 (DNAJB6) is also referred to as DnaJ's mammalian relative (MRJ). It is a cochaperone belonging to the heat shock protein HSP40 family [21]. We previously found that the expression of DNAJB6 was significantly upregulated in CRC tissues, and its overexpression was an independent predictor of poor outcomes in CRC patients [22]. KIAA1522 is a large protein-coding gene with an uncertain function [23]. It has been reported that the KIAA1522 overexpression has prognostic significance in several cancer types, including CRC [24–26]. Consistently, our earlier findings also suggested that KIAA1522 isa potentialprognostic marker in CRC.

In this study, we evaluated the expression of the three proteins and their impacts on the prognosis of CRC patients. Moreover, we constructed two prediction nomograms that integrated the three protein expression levels and TNM stage to predict the disease-free survival (DFS) and overall survival (OS) of patients with CRC after surgery.

Materials and methods

Patient & specimen characteristics

In total, 329 surgically resected CRC and morphologically normal operative margin tissues were collected between 2007 and 2011 at Cancer Hospital, CAMS/PUMC, Beijing in accordance with the provisions of the Helsinki Declaration. The inclusion criteria were as follows: (1) underwent radical resection of the primary tumor; (2) pathologically confirmed CRC; and (3) morphologically normal operative margin tissues. The exclusion criteria were as follows: (1) the presence of other malignant tumors; (2) incomplete follow-up data; (3) palliative surgery such as bypass surgery; (4) emergency operation such as bowel perforation or obstruction; and (5) neoadjuvant chemoradiotherapy. Clinical parameters including sex, age, pathologic T stage (pT), pathologic N stage (pN), pathologic M stage (pM), TNM stage, histological grade, tumor site, adjuvant chemotherapy therapy, vascular tumor thrombus, and nerve invasion were collected. TNM stage was classified based on the American Joint Committee on Cancer (AJCC) staging manual (seventh edition). The clinicopathologic parameters of all included patient are provided in Supplementary Table 1.This study was approved by the Ethics Committee/Institutional Review Board of National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College (NCC2020C-105). All patients signed informed consent forms.

The patients were divided into a training set (n = 230) and a validation set (n = 99), including 142 (61.7%) and 60 (60.6%) male patients in the training and validation sets, respectively. In total, 149 patients (45.3%) received adjuvant chemotherapy. Approximately 80% of these patients received oxaliplatin-containing regimens, such as mFOLFOX6 (oxaliplatin, fluorouracil, and leucovorin) or XELOX (oxaliplatin plus capecitabine). The baseline clinicopathological features of CRC in training and validation cohorts are summarized in Supplementary Table 2. For the entire cohort, 33.4% (110/329) of the patients suffered tumor recurrence, and 40.4% (133/329) died during follow-up. The median follow-up times were 62 months in the training cohort and 63.5 months in the validation cohort. Disease-free survival (DFS) and overall survival (OS) were primary study endpoints. DFS was defined as time from surgery to recurrence, distant metastasis or last follow-up, and OS was defined as time from surgery to death or last follow-up. The 1-, 3-, and 5-year OS rates were 95.7%, 73.0%, and 63.9%, respectively, and the 1-, 3-, and 5-year DFS rates were 90.8%, 68.3%, and 64.1%, respectively.

Tissue microarray and immunohistochemistry

Tissue microarrays (TMAs) were created using paired CRC and normal operative margin tissues, and IHC analysis was performed as previously described [22]. Each case in the TMAs included 2 representative normal mucosa regions and 3 representative tumor regions. The following antibodies were used: anti-p-mTOR antibody (Ser2448, #2976; Cell Signaling Technology, MA, USA); anti-DNAJB6 antibody (66587-1-Ig; Proteintech); and anti-KIAA1522 antibody (HPA032050; Sigma).

The intensity of the immunoreaction and the percentage of the staining area were used to calculate the levels of the three proteins. Staining intensity was graded as follows: 0 (negative), 1 (weak), 2 (moderate), and 3 (strong positive). The percentage of immunoreactive cells was assessed as 0 ($\leq$ 10%), 1 (11–25%), 2 (26–50%), 3 (51–75%), and 4 ($\geq$ 75%). The staining intensity and the percentage of the immunoreactive cells were multiplied to obtain the total score for each TMA dot [20]. According to the average expression scores of the specimens, the scores were rated as weak ($<$ 3) or strong ($\geq$ 3). The results were assessed blindly by two independent observers, and then discordant cases were reanalyzed jointly to reach an agreement. The expression level of the three proteins was stratified into low ($<$ 3) or high ($\geq$ 3) based on the IHC score.

Hierarchical clustering analysis

All raw score data were employed for clustering analysis. By using hierarchical clustering, tumors were categorized according to the relatedness of their immunostaining profile. The dendrogram on the top shows the clustering of the individual cases based on the degree of similarity of their immunohistochemical staining results. The longer the horizontal dendrogram arm is, the greater the difference in immunoprofiles between individual cases inside a cluster group. Heatmap generation and hierarchical clustering analysis (distance, standard Euclidean distance; criterion, complete-link) were performed using R 4.2.0 (http://www.r-project.org).

Nomogram construction and validation

The nomograms were constructed as follows. First, a univariate Cox proportional hazards model was used to assess the parameter's potential power in predicting the survival of the training cohort. Second, variables with *P* values less than 0.1 in univariate analysis were included in the multivariate Cox proportional hazards model. Finally, two nomograms containing the independent prognostic variables and TNM stage were created using R 4.2.0 (http://www.r-project.org) and the "rms" package. The best cutoffs for survival curve risk stratifications were determined by X-tile based on the composite scores of the nomograms [27]. Risk stratifications based on the nomogram models were evaluated using the Kaplan-Meier method.

The nomograms were subjected to internal validation in the training

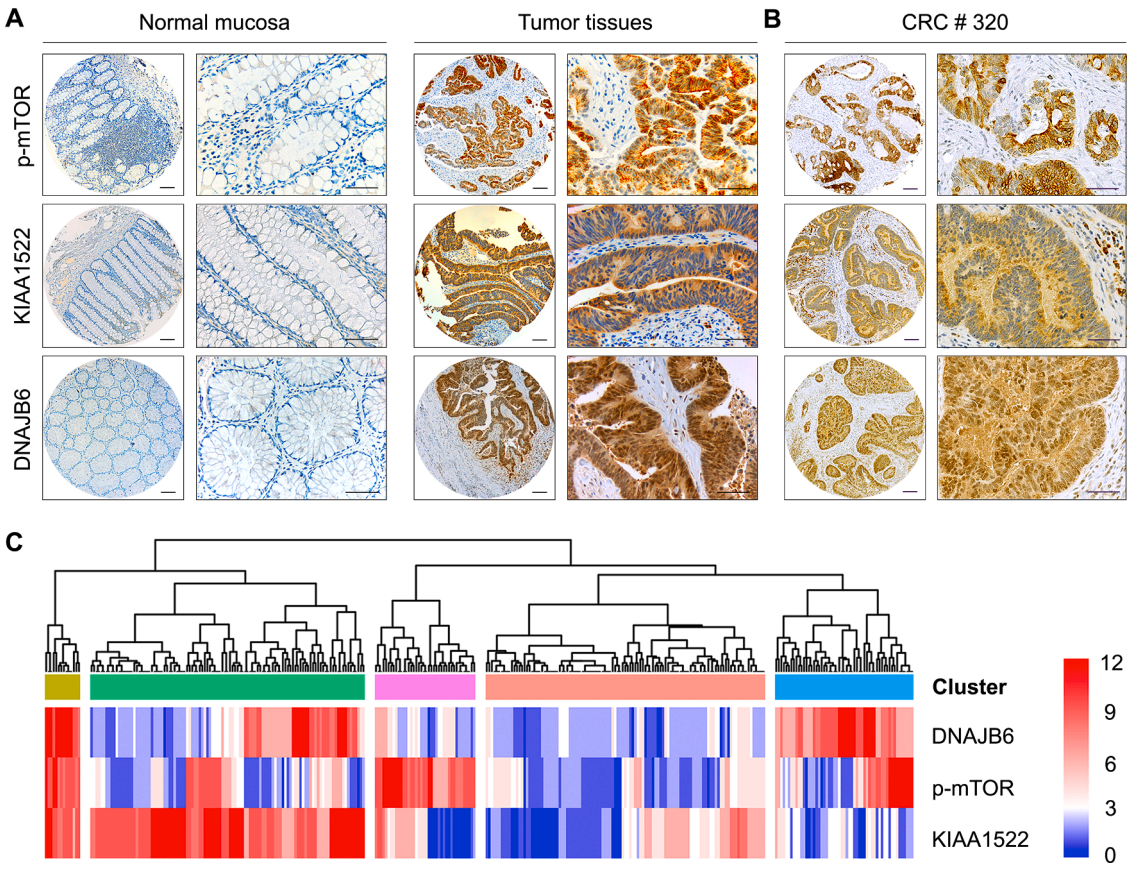

Fig. 1. MAP4, DNAJB6 and p-mTOR were overexpressed in CRC tissues.
A-B. Representative immunohistochemical results. **A**. Negative expression of MAP4, p-mTOR (Ser2448) and DNAJB6 in adjacent normal colorectal mucosa (left panel) and strong positive expression of these proteins in CRC tissues (right panel). **B**. Immunostaining results of the three proteins within a CRC specimen. The short- and long-scale bars represent 100 μm and 50 μm, respectively. **C.** Hierarchical cluster analysis of CRC TMA immunostaining results. Each column represents a single case, and each row represents a single marker. IHC scores range from 0 to 12 points, with a cutoff value of 3 points. Scores above 3 are shown in red, and scores below 3 are shown in blue. The shade of color indicates how different the IHC score was from the cutoff value.

cohort and external validation in the validation cohort. The validation of the nomogram was performed using the receiver operating characteristic (ROC) curves, the concordance index (C-index), calibration curves, and decision curve analysis (DCA). The C-index and area under the ROC curve (AUC) were used to evaluate the discriminative ability [28–31]. The C-index and AUC have values ranging from 0 to 1, where a value of 0 indicates an entirely inaccurate model and a value of 1 indicates a very accurate model. Generally, C-index and AUC values of more than 0.7 often indicate a reliable estimation [32]. In addition, the predicted survival probabilities by nomogram and actual survival probabilities were compared using a calibration plot. For an ideal nomogram, the predictive survival rates should fall on the 45-degree diagonal line [33]. DCA was conducted to assess the clinical practical value of the nomogram models by quantifying the net benefit at different threshold probabilities [34]. The net benefit is defined as that true positive minus false positive [35]. In brief, the curves of all patients died and none of the patients died were plotted as two references. DCA calculates the clinical benefit compared with the reference lines. The higher the net benefit, the more practical and effective the prediction model is in clinical practice.

The flow chart of the construction and validation of the nomogram models is shown in Supplementary Fig. 1.

Statistical analysis

Statistical analyses were performed with SPSS 26.0 (IBM Corp, Armonk, NY, USA). Violin plots and histograms were generated with GraphPad Prism software. Paired Student's t test was used to assess the difference in protein expression between CRC and adjacent normal tissues. The Pearson chi-square test and Fisher's exact test were performed to identify associations between protein levels and clinicopathological characteristics. Survival curves for protein biomarkers were generated using the Kaplan-Meier method and were compared using the log-rank test. Univariable and multivariable Cox proportional hazards models were used to analyze DFS and OS, in which hazard ratios (HRs) and their 95% confidence intervals (CIs) were calculated. Statistical significance was defined as a P value of less than 0.05.

Results

KA1522, DNAJB6, and p-mTOR were upregulated in CRC tissues

The protein expression levels of KIAA1522, DNAJB6, and p-mTOR (Ser2448, an activated form of mTOR) were examined via IHC on TMAs containing 329 CRC specimens. Immunohistochemistry data of each patient are provided in Supplementary Table 1 and representative immunostaining images and hierarchical cluster analysis results are shown in Fig. 1. KIAA1522 protein was mainly expressed in the cytoplasm, p-mTOR was expressed in the cytomembrane and cytoplasm, and DNAJB6 staining was observed in both the cytoplasm and nucleus. Significantly high KIAA1522, DNAJB6, and p-mTOR expression levels were detected in 69.3% (228/329), 55.9% (184/329), and 56.2% (185/329) of the CRC tissues, respectively, but their expression was generally absent in normal colorectal mucosa (each $P < 0.001$, Supplementary Table 3).

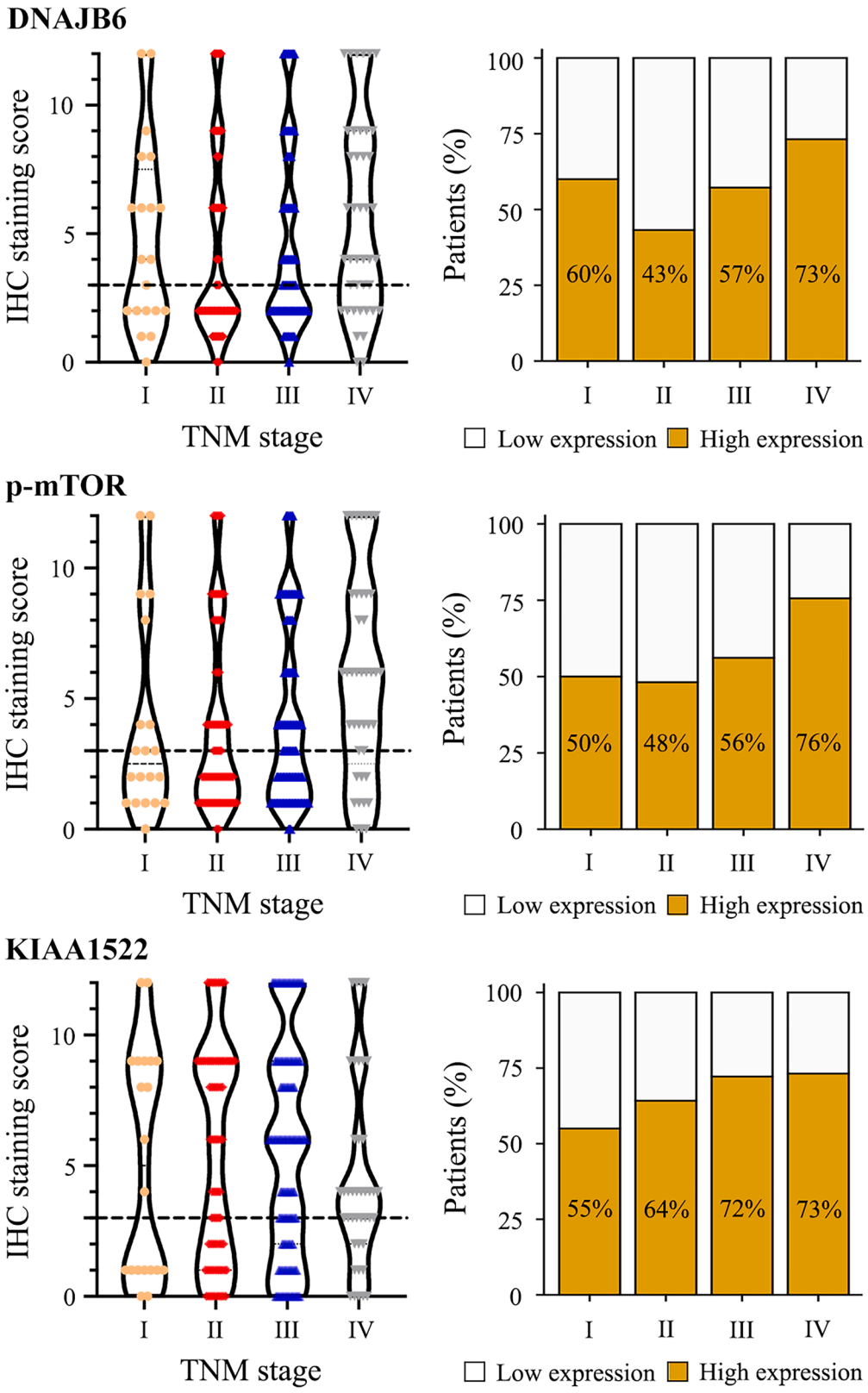

Fig. 2. DNAJB6, p-mTOR and KIAA1522 expression in different stages of CRC. Left panel: Violin plots of IHC scores for CRC patients at different cancer stages. Dashed line indicates the cutoff value of IHC score (i.e., 3). Based on IHC score, protein expression was divided into low (< 3) or high (≥ 3). Right panel: Histogram of percentage of protein marker expression in CRC patients based on TNM staging.

Table 1
Relationship between protein expression and clinicopathologic parameters

Variables	Total No.	KIAA1522 expression (%)			DNAJB6 expression (%)			p-mTOR expression (%)		
		high	low	*p*	high	low	*p*	high	low	*p*
Age (years)				0.053			**0.013**			0.415
<60	179	116 (50.9%)	63 (62.4%)		89 (48.4%)	90 (62.1%)		97 (52.4%)	82 (56.9%)	
≥60	150	112 (49.1%)	38 (37.6%)		95 (51.6%)	55 (37.9%)		88 (47.6%)	62 (43.1%)	
Sex				0.463			0.169			0.314
female	127	91 (39.9%)	36 (35.6%)		65 (35.3%)	62 (42.8%)		67 (36.2%)	60 (41.7%)	
male	202	137 (60.1%)	65 (64.4%)		119 (64.7%)	83 (57.2%)		118 (63.8%)	84 (58.3%)	
TNM Stage				0.268			**0.015**			**0.033**
I	20	11 (4.8%)	9 (8.9%)		12 (6.5%)	8 (5.5%)		10 (5.4%)	10 (6.9%)	
II	81	52 (22.8%)	29 (28.7%)		35 (19.0%)	46 (31.7%)		39 (21.1%)	42 (29.2%)	
III	187	135 (59.2%)	52 (51.5%)		107 (58.2%)	80 (55.2%)		105 (56.8%)	82 (56.9%)	
IV	41	30 (13.2%)	11 (10.9%)		30 (16.3%)	11 (7.6%)		31 (16.8%)	10 (6.9%)	
pT				0.482*			0.887*			0.164*
T1	2	1 (0.4%)	1 (1.0%)		1 (0.5%)	1 (0.7%)		0 (0%)	2 (1.4%)	
T2	33	20 (8.8%)	13 (12.9%)		19 (10.3%)	14 (9.7%)		17 (9.2%)	16 (11.1%)	
T3	211	147 (64.5%)	64 (63.4%)		115 (62.5%)	96 (66.2%)		115 (62.2%)	96 (66.7%)	
T4	83	60 (26.3%)	23 (22.8%)		49 (26.6%)	34 (23.4%)		53 (28.6%)	30 (20.8%)	
pN				0.068			0.076			**0.036**
N0	132	84 (36.8%)	48 (47.5%)		66 (35.9%)	66 (45.5%)		65 (35.1%)	67 (46.5%)	
N+	197	144 (63.2%)	53 (52.5%)		118 (64.1%)	79 (54.5%)		120 (64.9%)	77 (53.5%)	
pM				0.771			0.501			**0.005**
M0	305	212 (93.0%)	93 (92.1%)		169 (91.8%)	136 (93.8%)		165 (89.2%)	140 (97.2%)	
M1	24	16 (7.0%)	8 (7.9%)		15 (8.2%)	9 (6.2%)		20 (10.8%)	4 (2.8%)	
Histological grade				0.921			0.132			0.807
G1	56	40 (17.5%)	16 (15.8%)		38 (20.7%)	18 (12.4%)		32 (17.3%)	24 (16.7%)	
G2	223	154 (67.5%)	69 (68.3%)		118 (64.1%)	105 (72.4%)		127 (68.6%)	96 (66.7%)	
G3	50	34 (14.9%)	16 (15.8%)		28 (15.2%)	22 (15.2%)		26 (14.1%)	24 (16.7%)	
Initial tumor site				0.110			0.860			0.312
Colon	122	91 (39.9%)	31 (30.7%)		69 (37.5%)	53 (36.6%)		73 (39.5%)	49 (34.0%)	
Rectum	207	137 (60.1%)	70 (69.3%)		115 (62.5%)	92 (63.4%)		112 (60.5%)	95 (66.0%)	
Vascular tumor thrombus				0.221			0.061			0.608
Yes	29	23 (10.1%)	6 (5.9%)		21 (11.4%)	8 (5.5%)		15 (8.1%)	14 (9.7%)	
No	300	205 (89.9%)	95 (94.1%)		163 (88.6%)	137 (94.5%)		170 (91.9%)	130 (90.3%)	
Nerve invasion				1.000*			1.000*			1.000*
Yes	1	1 (0.4%)	0 (0%)		1 (0.5%)	0 (0.0%)		0 (0%)	1 (0.7%)	
No	328	227 (99.6%)	101 (100%)		183 (99.5%)	145 (100.0%)		185 (100%)	143 (99.3%)	
Adjuvant Chemotherapy therapy				0.763			0.234			0.861
Presence	149	102 (44.7%)	47 (46.5%)		78 (42.4%)	71 (49.0%)		83 (44.9%)	66 (45.8%)	
Absence	180	126 (55.3%)	54 (53.5%)		106 (57.6%)	74 (51.0%)		102 (55.1%)	78 (54.2%)	

pT: pathologic T stage; pN, pathologic N stage, pM, pathologic M stage; * Fisher's exact test.

Additionally, the percentage of patients with high KIAA1522, DNAJB6, and p-mTOR expression tended to increase as the disease progressed from TNM stage I to stage IV (Fig 2).

Statistical analysis showed that aberrant mTOR activation was positively associated with pN ($P = 0.036$), pM ($P = 0.005$), and TNM stage ($P = 0.033$); DNAJB6 overexpression was significantly correlated with patient age ($P = 0.013$) and TNM stage ($P = 0.015$). There were no significant links between abnormally expressed KIAA1522 and any clinical characteristics in this study (Table 1).

Prognostic significance of the three protein biomarkers in the training cohort

Next, we evaluated the impact of the overexpression of the three proteins on the clinical outcomes of patients with CRC in the training set. As a single biomarker, a high abundance of KIAA1522, DNAJB6, or p-mTOR protein was strongly correlated with poorer DFS and OS in CRC patients (each $P < 0.01$, Fig. 3). Univariate analysis of the training cohort revealed that pN, pM, initial tumor site, vascular tumor thrombus, nerve invasion, and the levels of the three proteins were linked to DFS in CRC patients (Supplementary Table 4). Multivariate Cox regression analysis revealed that DNAJB6 overexpression (HR = 1.73, 95% CI: 1.05-2.84, $P = 0.031$), mTOR hyperactivation (HR = 2.34, 95% CI: 1.34-4.10, $P = 0.003$), and KIAA1522 overexpression (HR = 1.98, 95% CI: 1.09-3.60, $P = 0.025$) were all independent risk factors for unfavorable DFS in CRC patients (Supplementary Table 4, Fig. 4).

Univariate and multivariate Cox regression models were used to assess each prognostic factor for the OS of CRC patients in the training cohort. As a result, pN, histological grade, and the expression of the three proteins were all substantially associated with OS in univariate regression analysis (all $P < 0.1$, Supplementary Table 5), and DNAJB6 overexpression (HR = 1.92, 95% CI: 1.22-3.05, $P = 0.005$), mTOR hyperactivation (HR = 1.71, 95% CI: 1.07-2.73, $P = 0.024$), and KIAA1522 overexpression (HR = 1.74, 95% CI: 1.04-2.91, $P = 0.033$) were identified as independent factors for inferior OS rates in multivariate Cox regression analyses (Supplementary Table 5, Fig. 4).

Construction of a protein-associated prognostic models using the training cohort

Given that the prognosis of CRC patients can rarely able be accurately predicted by any single biomarker or clinical feature, we combined the KIAA1522, DNAJB6, and p-mTOR protein markers with several well-defined clinical prognostic factors (i.e., pT, pN, and pM) to create comprehensive prognostic nomograms for evaluating the OS and DFS probabilities in CRC patients (Fig. 5). In these nomograms, the length of each variable line reflects its contribution to prognosis. For instance, our nomograms showed that pT had the most prominent impact on both DFS and OS in CRC patients among the included clinical parameters. Of note, the expression levels of the three protein biomarkers exhibited a remarkable influence on the outcome of CRC patients, especially on the DFS of patients. Collectively, the KIAA1522 expression level contributed the most to predicting the DFS of CRC patients among all the parameters. In addition, compared with pN and pM,

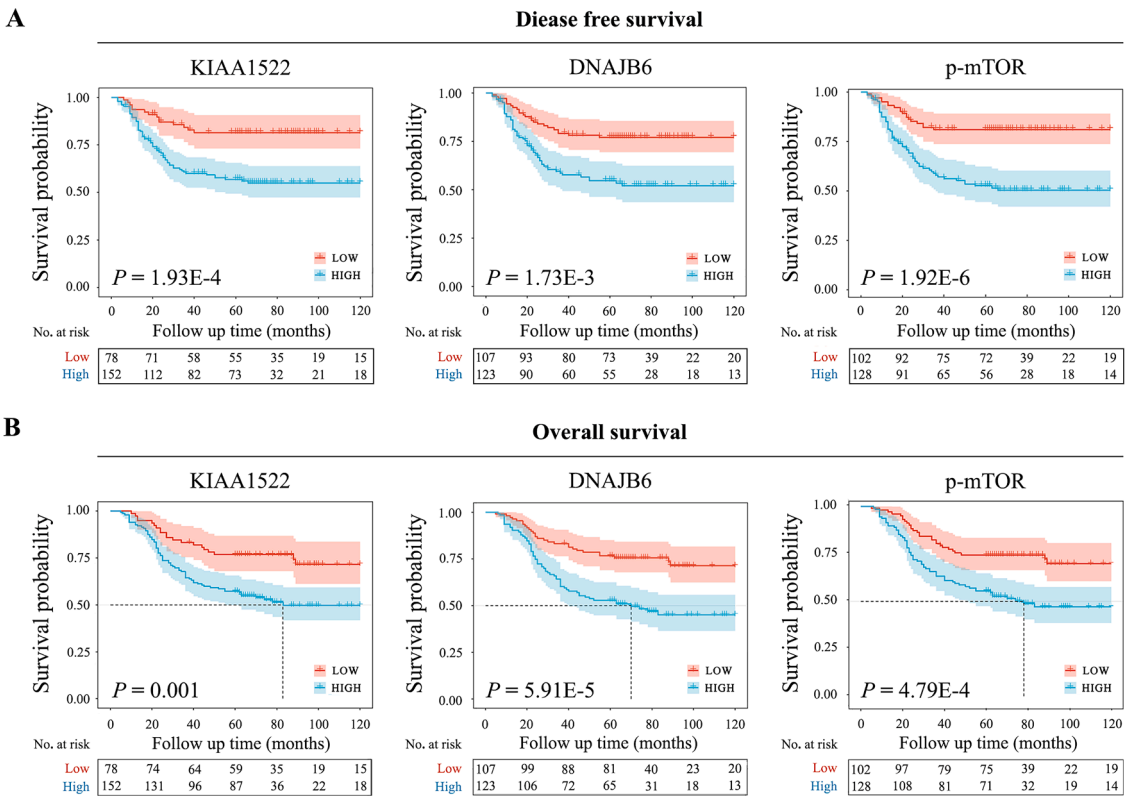

Fig. 3. The relationship between protein expression and survival of CRC patients. Kaplan-Meier survival analysis of disease-free survival (**A**) or overall survival (**B**) of CRC patients according to the indicated protein expression.

the expression levels of the three proteins were better predictors of both DFS and OS of CRC patients (Fig. 5).

To use the nomograms to predict the prognosis of an individual CRC patient, first determine the score for every variable based on the value on the topmost point row corresponding to its parameter. For example, the score of T1 in the DFS nomogram is approximately 32.5; the score of a low level of p-mTOR in the OS nomogram is 35. Then, the scores of all variables can be summed to obtain the total score. The probability of 1-, 3-, and 5-year OS and DFS for the patient can be predicted based on the values on the OS or DFS row corresponding to the total score. A lower total score is associated with a worse outcome. For example, a patient with high DNAJB6, p-mTOR and KIAA1522 expression, pT3, pN0, and pM0 CRC would have a total score of 89 based on the OS nomogram (0 points for high expression of the three proteins, 45 points for pT3, 29 points for pN0 and 15 points for pM0). For this patient, the predicted 1-, 3-, and 5-year OS rates are approximately 95.0%, 72.5%, and 58.0%, respectively.

CRC patients can be classified into low- and high-risk groups based on their nomogram scores. In brief, the total scores of the nomograms were used to divide the patients into two groups using X-tile. The threshold value was 242.1 points in the DFS nomogram and 108.4 points in the OS nomogram. Patients with points higher than the threshold were classified into the low-risk group. Survival analysis confirmed that the DFS and OS probabilities in the high-risk group were significantly lower than that in the low-risk group, suggesting that these nomograms may be used for risk stratification in CRC patients (Fig. 6).

Validation of the protein-associated prognostic model

Subsequently, we evaluated the performance of the prognostic models using the calibration plot, c-index, AUC values, and DCA curves in both the training and validation cohorts. First, corresponding calibration curves were drawn, and the results showed good agreement between the 1-, 3-, and 5-year DFS and OS probabilities predicted by the nomogram models and the actual outcomes of CRC patients in the training and validation cohorts, indicating that the prediction models have high accuracy (Fig. 7).

Second, ROC curves were plotted to assess the predictive sensitivities and specificities of the nomogram prediction models. The AUC values for the prediction models were approximately 0.70 in the training and validation cohorts (Fig. 8A). Additionally, the AUC values of the nomogram models were higher than those of the 3-protein markers or TNM stage in the entire cohort, suggesting the superior prediction performance of our nomogram models (Supplementary Fig. 2). Likewise, the C-index was also calculated to measure the discrimination ability of the nomogram models. In the training cohort, the C-indices for the prediction of DFS and OS were 0.720±0.055 and 0.692±0.054, respectively. In the validation cohort, the C-indices for the prediction of DFS and OS were 0.635±0.097 and 0.647±0.087, respectively. Overall, these data corroborated that the sensitivities and specificities of our nomogram models were relatively high.

Furthermore, DCA was performed to assess the clinical utility of the nomogram prediction models by quantifying the net benefits at different threshold probabilities. Theoretically, the higher the net benefit is, the more practical and effective the prediction model is in clinical practice. As a result, for almost all of the threshold probabilities in both the training and validation cohorts, higher net benefits were observed with the prediction model than in two extreme cases (i.e., all patients died or relapsed or none of the patients died or relapsed) (Fig. 8B). Notably, both the DFS and OS nomograms had a higher net benefit than TNM staging when the threshold probability was $>$ 10% in the training cohort. Collectively, these data indicate that our nomogram prediction models have high clinical utility.

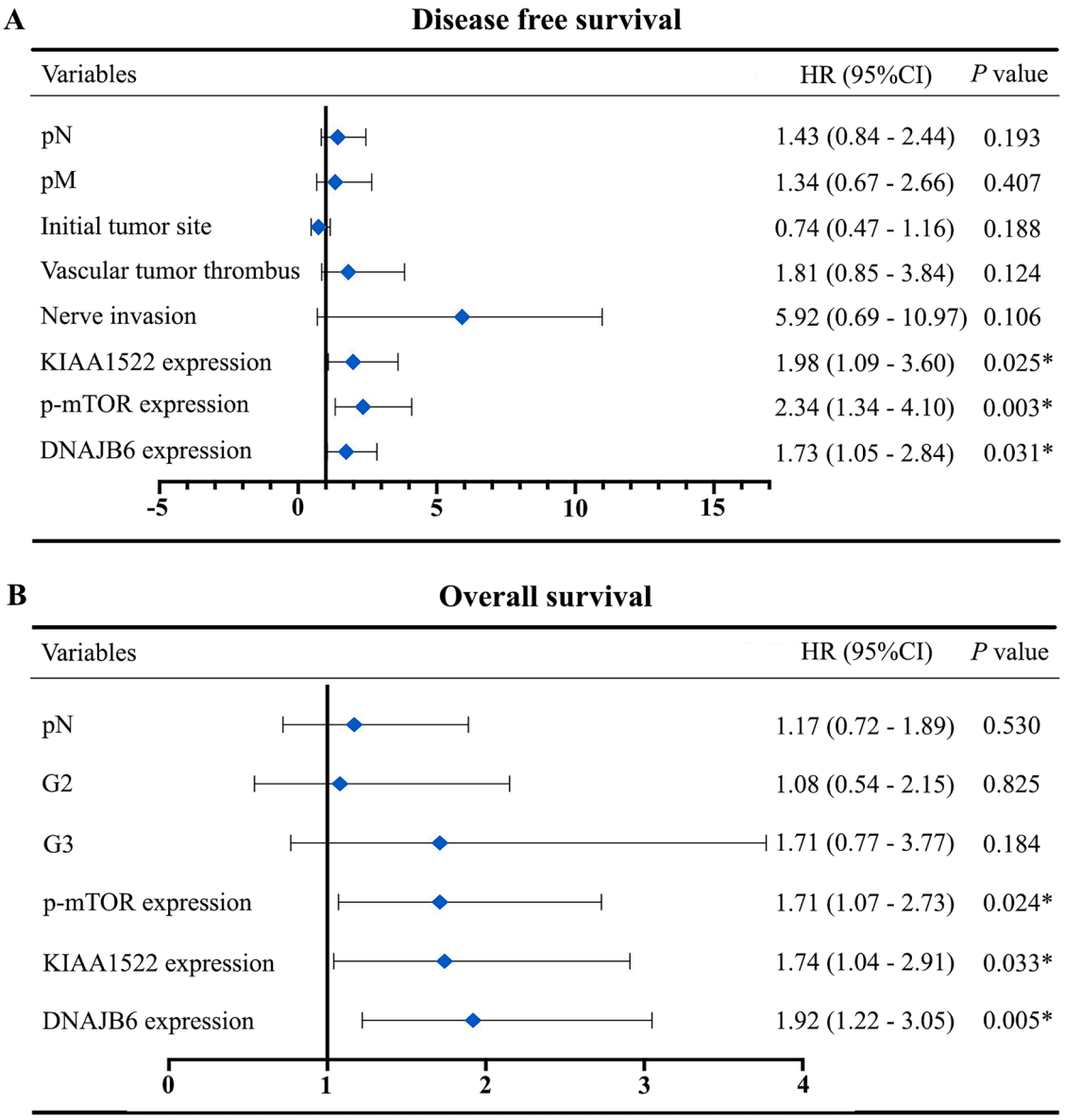

Fig. 4. Forest plots of multivariate Cox regression analysis for disease-free survival and overall survival of patients with CRC. A. Disease-free survival; **B.** Overall survival. HR: hazard ratio; CI: confidence interval.

Discussion

CRC is a heterogeneous malignancy with a high risk of recurrence and death. There are few effective therapeutic approaches for recurrent and metastatic CRC to date. Therefore, it is imperative to develop effective prognostic prediction models to identify high-risk patients and carry out early intervention and individualized treatment. To this end, we combined three protein biomarkers and the TNM stage to establish nomograms that potentially predict the risk of death and recurrence in individual CRC patients in this study.

Clinical pathological characteristics such as lymph node count and histological grading are frequently used in clinical settings for risk assessment and therapy decision-making for CRC [20]. However, CRC is a highly heterogeneous disease with varying genetic backgrounds, and it is challenging to predict patients' outcomes, even among tumors with identical histopathological features. As a result, the addition of molecular biomarkers may improve the accuracy of patient prognosis prediction. Molecular biology and bioinformatics techniques have been widely employed to uncover biomarkers, such as DNA and RNA, to assess the prognosis of patients with carcinoma [36–38]. In comparison, protein biomarker identification using IHC, which is widely used in clinical diagnosis, has been discovered to be a rapid, cost-effective, and accurate technique for tumor molecular profiling.

Previous works have developed nomograms integrating clinical data and protein biomarkers to predict the risk of recurrence or mortality for patients with CRC. However, most of these studies only focused on tumor suppressors with reduced expression in tumor tissues, such as E-cadherin, CD44, and CSN2, and most studies focused on the prognostic role of a single protein marker [39,40]. In this study, we found that DNAJB6, p-mTOR, and KIAA1522 were upregulated in CRC tissues. High expression levels of the three proteins were statistically associated with unfavorable OS and DFS in CRC patients. Considering that clinical risk factors, such as TNM stage, remain significant predictors of survival in CRC, we integrated TNM stage with our three protein biomarkers to develop efficient models for predicting survival to avoid omitting valuable variables.

The protein biomarkers analyzed in this study were carefully chosen based on our previous research findings. For example, we found that aberrant expression of DNAJB6 was significantly correlated with poor outcomes of CRC patients [22]. More importantly, a functional study

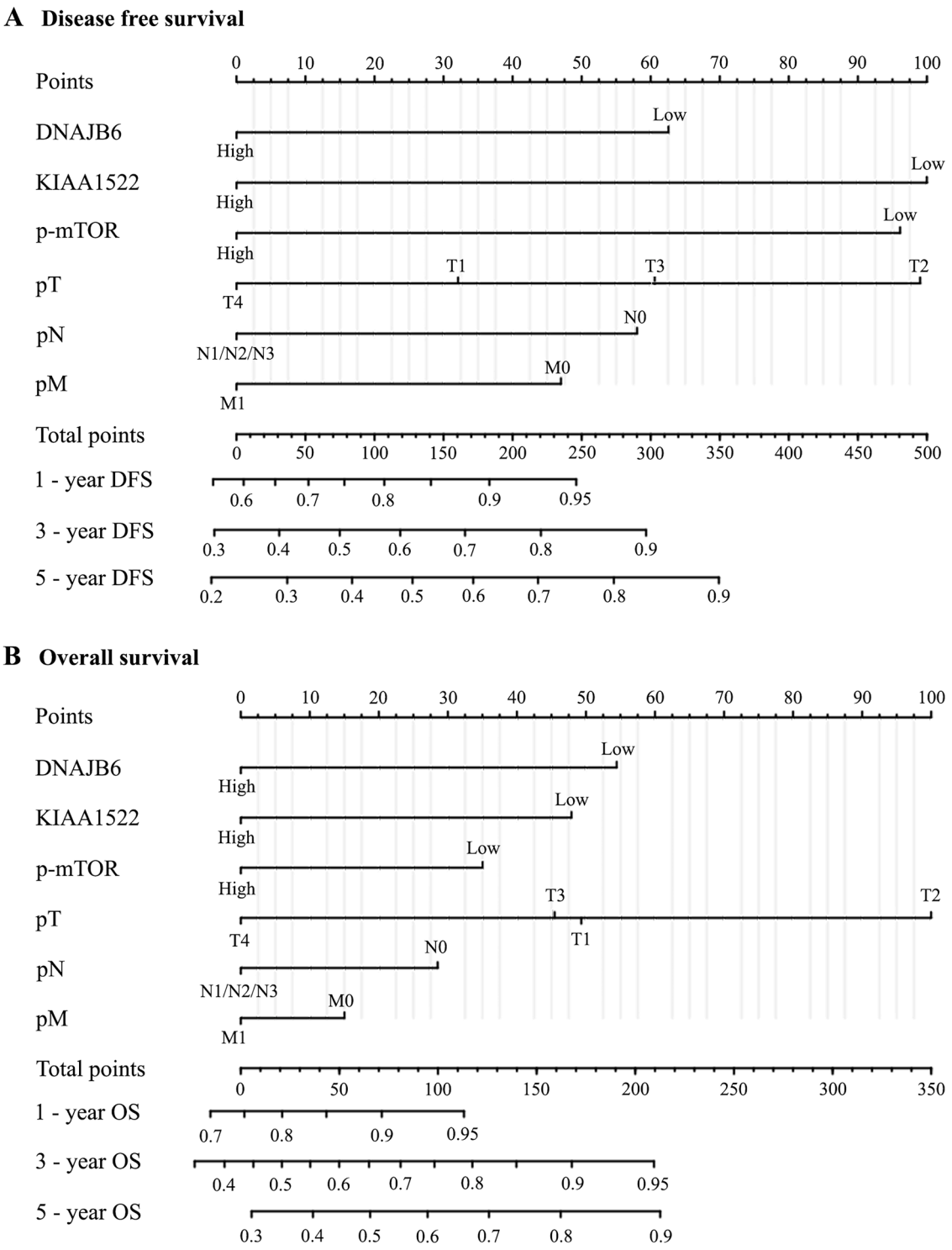

Fig. 5. Nomograms for predicting the prognosis of patients with CRC. Nomograms for predicting the disease-free survival (**A**) and overall survival (**B**) of CRC patients were created by integrating the indicated protein expression and several pivotal clinical prognostic factors. The score of each risk factor can be determined by drawing a vertical line straight upward from the factor's corresponding parameter to the points axis. Then, add the scores of all risk factors together, and draw a straight line down from the total points axis to the OS or DFS axis to obtain the survival probabilities of CRC patients 1, 3, and 5 years postoperation.

revealed that DNAJB6 overexpression enhanced the metastatic potential of CRC cells, which further suggested that its high expression is a crucial risk factor in the prognostic assessment of CRC [22]. In the current study, we validated the prognostic significance of the aberrant expression of DNAJB6b in a larger CRC sample size. Our data showed that DNAJB6 was a promising biomarker for dividing CRC patients into subgroups with varying risks of death and recurrence. However, its function and underlying mechanism in CRC are still not well understood and need further study.

Although overactivation of mTOR is correlated with a poor prognosis in most cancer types, the prognostic significance of mTOR activation in CRC is still controversial [41,42]. Our earlier data revealed that p-mTOR was a potential prognostic marker in CRC [20]. In this study, high p-mTOR activity was also determined to be an independent risk factor

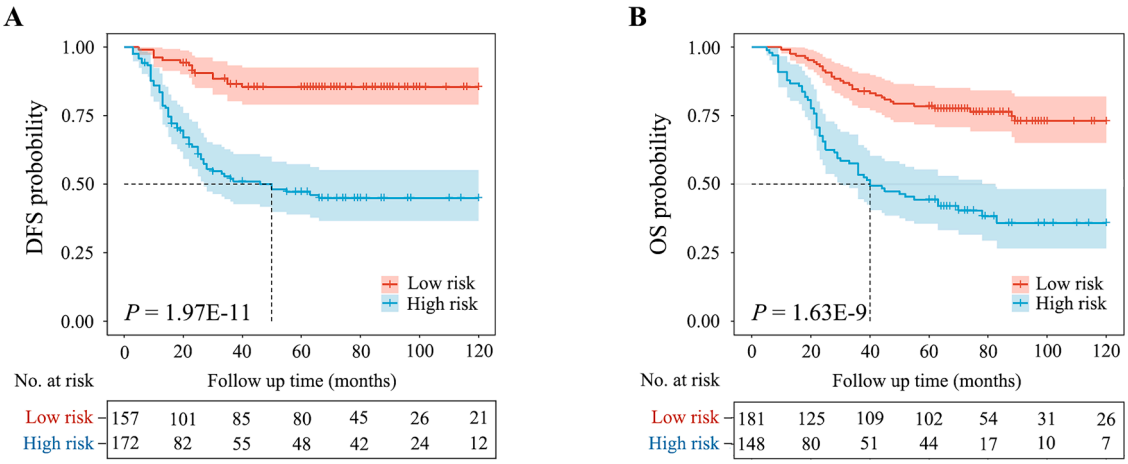

Fig. 6. Nomogram-based risk stratification. CRC patients were divided into low- and high-risk subgroups by the nomogram score, and Kaplan-Meier survival analysis was performed to verify the clinical significance of the nomogram models. A. Disease-free survival; **B.** Overall survival.

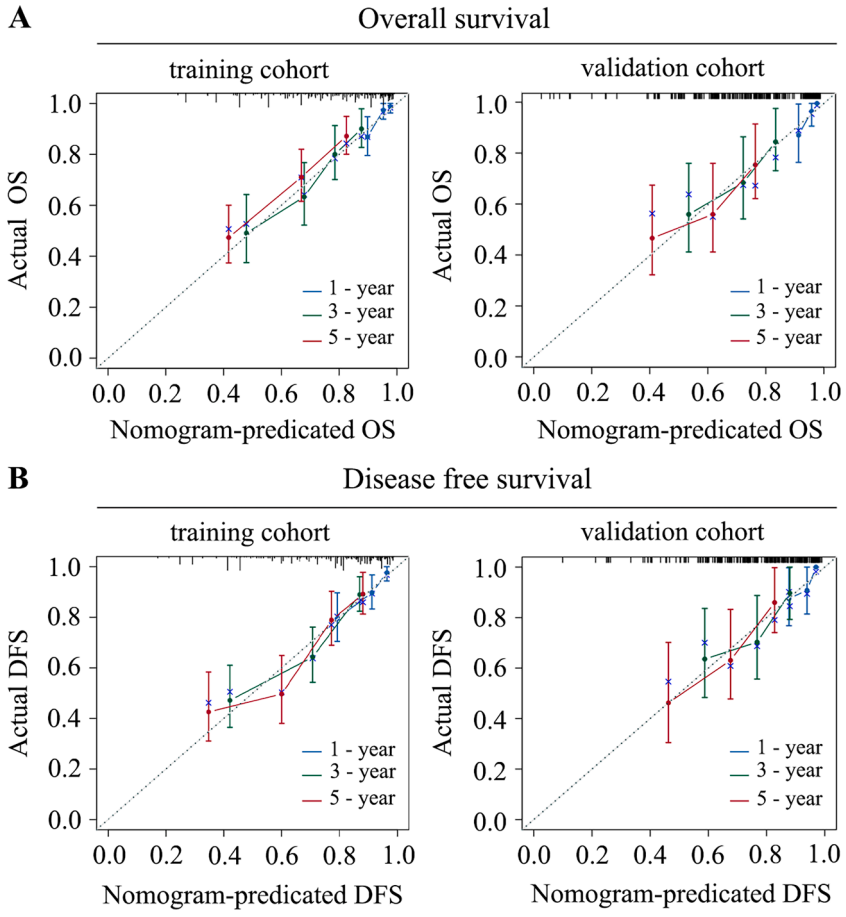

Fig. 7. Calibration curves for predicting the survival of CRC patients. A. Overall survival. **B.** Disease-free survival. The 45-degree dotted line represents an ideal nomogram, and the solid line represents our nomogram. If the predicted survival probability is on the 45-degree diagonal, it means the prediction is accurate.

for CRC survival. Consistent with our findings, a study reported that mTOR was aberrantly activated in 73.8% of cases, and a high level of p-mTOR was an independent adverse prognostic indicator in CRC patients [43]. However, in other studies, mTOR activation was not significantly linked to the survival of CRC patients [42,44]. Therefore, more research may be required to fully understand the link between mTOR hyperactivation and CRC prognosis. Notably, a series of studies have shown that mTOR signaling can regulate the proliferation, cell cycle progression, EMT, motility and invasion of CRC cells [45–47], suggesting that abnormal activation of mTOR signaling predicts the unfavorable outcomes of CRC patients.

We previously found that KIAA1522 was overexpressed in non-small cell lung cancer (NSCLC) and was associated with a poor prognosis in patients with NSCLC [25]. In addition, one study suggested that KIAA1522 might be a novel prognostic marker of hepatocellular carcinoma [26]. Recently, Yi et al. reported that KIAA1522 was an

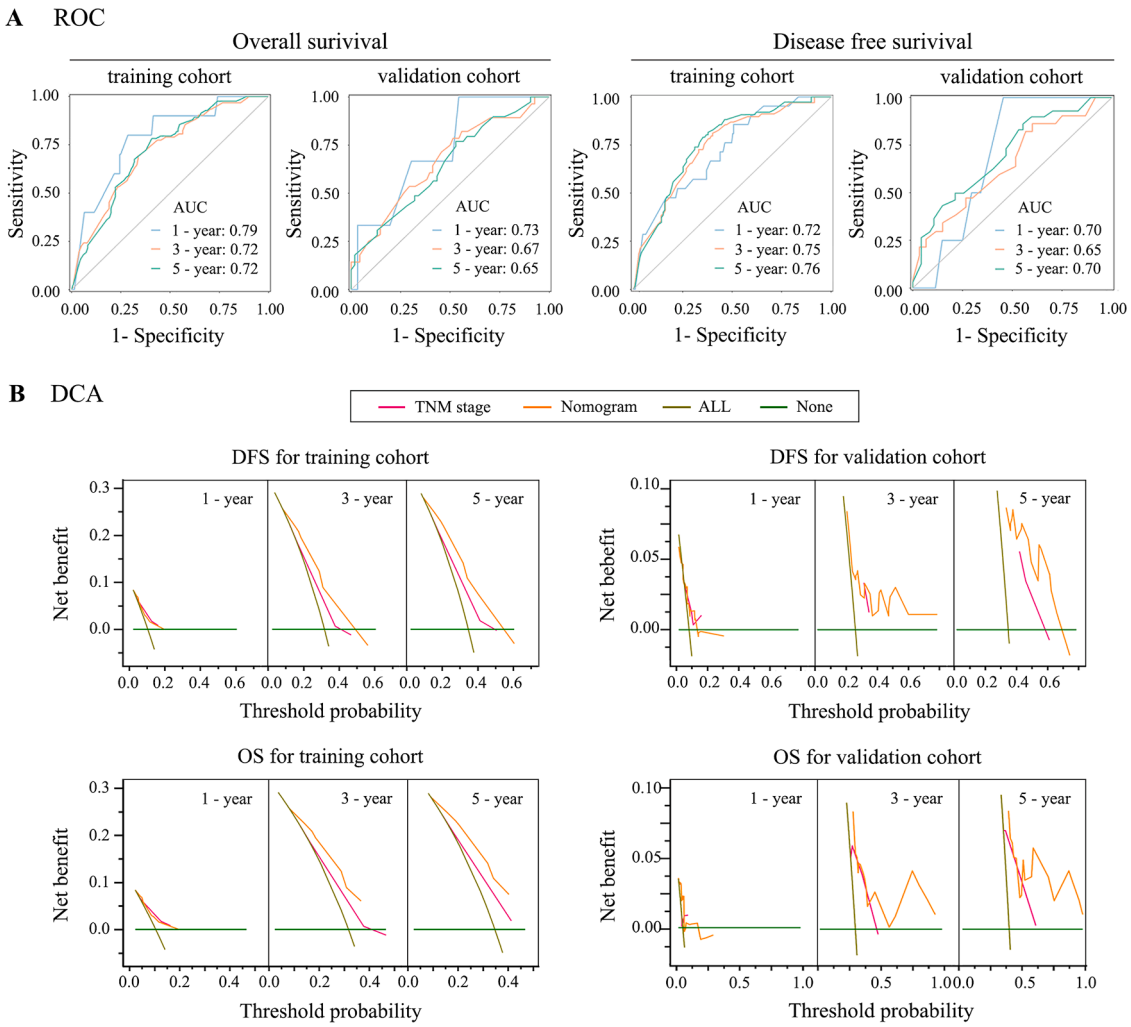

Fig. 8. Validation of the prognostic nomograms using ROC curves and decision curve analysis. A. ROC curves; **B.** Decision curve analysis. AUC, area under the curve. All, all patients died or relapsed; None, no patients died or relapsed.

independent prognostic biomarker of patients with colorectal carcinoma liver metastasis. Furthermore, functional studies in vitro and in vivo suggested that KIAA1522 might play an essential role in the progression of CRC [24]. Our current data confirmed that KIAA1522 was overexpressed in CRC tissues and that its aberrant expression was an independent prognostic marker in CRC. The above findings also suggested that KIAA1522 might be a potential therapeutic target in CRC. Functional studies revealed that high expression of KIAA1522 promoted the proliferation, survival, migration and invasion of CRC cells [48]. Whether KIAA1522 has other functional roles in CRC cells, such as regulating the cell cycle and EMT, remains to be further investigated.

Since CRC has high intratumor heterogeneity, it is difficult to accurately predict the outcome of CRC patients with a single biomarker [49, 50]. In the present study, we established novel prognostic prediction nomograms by combining the three protein markers and TNM stage. More importantly, we found that the combined nomograms were more accurate and practical than the three-protein biomarker and TNM stage alone in predicting the prognosis of CRC patients. Performance evaluations indicated that our nomograms had satisfactory discriminability, sensitivity, specificity. In addition, the clinical benefits of our models increased with the length of the follow-up period, suggesting that the models may be helpful for the long-term management of CRC patients. By using prediction models, the risk of recurrence or death of CRC patients can be more accurately evaluated, which may help clinicians select personalized and precise treatment plans for individual patients. Notably, p-mTOR, KIAA1522, and DNAJB6 may be potential therapeutic targets for CRC treatment. According to our prediction models, CRC patients with high expression levels of the three proteins have poorer outcomes and might benefit from the corresponding targeted therapy in the future.

In conclusion, we established novel prognostic prediction models that combined three protein biomarkers with TNM stage for the molecular classification of CRC patients in the present study. The prediction models could provide valuable prognostic information, including the risk of mortality and recurrence in individual CRC patients, thus facilitating the individualized and precise treatment of CRC. Nevertheless, the performance of the models needs to be validated by multicenter prospective studies in the future. In addition, the functional roles of these proteins and related molecular mechanisms in CRC cells should be further clarified.

Availability of data and materials

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Declarations of interest

None

CRediT authorship contribution statement

Yu-Juan Jiang: Investigation, Formal analysis, Methodology, Writing – original draft. **Tong-Tong Zhang:** Investigation, Formal analysis. **Yi-Qing Zhu:** Formal analysis. **Hong-Qing Cai:** Methodology. **Chen Chang:** Resources. **Jia-Jie Hao:** Methodology. **Yan Cai:** Resources. **Ming-Rong Wang:** Supervision. **Jian-Wei Liang:** Funding acquisition, Resources, Writing – review & editing. **Yu Zhang:** Funding acquisition, Conceptualization, Project administration, Data curation, Writing – review & editing.

Declaration of Competing Interest

The authors declare no potential conflicts of interest regarding this study.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (82073093, 81441070) and Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences (CIFMS, 2021-I2M-1-018).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tranon.2022.101609.

References

[1] H Sung, J Ferlay, R L Siegel, et al., Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J. Clin. 71 (3) (2021) 209–249, https://doi.org/10.3322/caac.21660.
[2] R Zheng, S Zhang, H Zeng, et al., Cancer incidence and mortality in China, 2016, J. Natl. Cancer Cent. (2022), https://doi.org/10.1016/j.jncc.2022.02.002 [2022-03-20], https://www.sciencedirect.com/science/article/pii/S2667005422000047.
[3] P Shu, J Wu, Y Tong, et al., Gene pair based prognostic signature for colorectal colon cancer, Medicine 97 (42) (2018) e12788, https://doi.org/10.1097/MD.0000000000012788.
[4] W Chang, X Gao, Y Han, et al., Gene expression profiling-derived immunohistochemistry signature with high prognostic value in colorectal carcinoma, Gut 63 (9) (2014) 1457–1467, https://doi.org/10.1136/gutjnl-2013-305475.
[5] J M Inadomi, Screening for colorectal neoplasia, N. Engl. J. Med. 376 (2) (2017) 149–156, https://doi.org/10.1056/NEJMcp1512286.
[6] E Koncina, S Haan, S Rauh, et al., Prognostic and predictive molecular biomarkers for colorectal cancer: updates and challenges, Cancers 12 (2) (2020) E319, https://doi.org/10.3390/cancers12020319.
[7] P.W. Laird, The power and the promise of DNA methylation markers, Nat. Rev. Cancer 3 (4) (2003) 253–266, https://doi.org/10.1038/nrc1045.
[8] M Zamani, S V Hosseini, P. Mokarram, Epigenetic biomarkers in colorectal cancer: premises and prospects, Biomarkers 23 (2) (2018) 105–114, https://doi.org/10.1080/1354750X.2016.1252961.
[9] S Kopetz, J Tabernero, R Rosenberg, et al., Genomic classifier ColoPrint predicts recurrence in stage II colorectal cancer patients more accurately than clinical factors, Oncologist 20 (2) (2015) 127–133, https://doi.org/10.1634/theoncologist.2014-0325.
[10] S Surinova, L Radová, M Choi, et al., Non-invasive prognostic protein biomarker signatures associated with colorectal cancer, EMBO Mol. Med. 7 (9) (2015) 1153–1165, https://doi.org/10.15252/emmm.201404874.
[11] F Pagès, B Mlecnik, F Marliot, et al., International validation of the consensus immunoscore for the classification of colon cancer: a prognostic and accuracy study, Lancet 391 (10135) (2018) 2128–2139, https://doi.org/10.1016/S0140-6736(18)30789-X.
[12] D Roy, A Lucci, M Ignatiadis, et al., Cell-free circulating tumor DNA profiling in cancer management, Trends Mol. Med. 27 (10) (2021) 1014–1015, https://doi.org/10.1016/j.molmed.2021.07.001.
[13] D Roy, M. Tiirikainen, Diagnostic power of DNA methylation classifiers for early detection of cancer, Trends Cancer 6 (2) (2020) 78–81, https://doi.org/10.1016/j.trecan.2019.12.006.
[14] A Li, D H Yang, Application of immunohistochemistry in basic and clinical studies, Immune Mediat. Cancer (2020) 43–55, https://doi.org/10.1007/978-1-0716-0247-8_4.
[15] C T Keith, S.L. Schreiber, PIK-related kinases: DNA repair, recombination, and cell cycle checkpoints, Science 270 (5233) (1995) 50–51, https://doi.org/10.1126/science.270.5233.50.
[16] Z Moafian, A Maghrouni, A Soltani, et al., Cross-talk between non-coding RNAs and PI3K/AKT/mTOR pathway in colorectal cancer, Mol. Biol. Rep. 48 (5) (2021) 4797–4811, https://doi.org/10.1007/s11033-021-06458-y.
[17] P du Rusquec, C Blonz, J S Frenel, et al., Targeting the PI3K/Akt/mTOR pathway in estrogen-receptor positive HER2 negative advanced breast cancer, Ther. Adv. Med. Oncol. 12 (2020), 1758835920940939, https://doi.org/10.1177/1758835920940939.
[18] H Li, J Zeng, K. Shen, PI3K/AKT/mTOR signaling pathway as a therapeutic target for ovarian cancer, Arch. Gynecol. Obstet. 290 (6) (2014) 1067–1078, https://doi.org/10.1007/s00404-014-3377-3.
[19] Y Dobashi, Y Watanabe, C Miwa, et al., Mammalian target of rapamycin: a central node of complex signaling cascades, Int. J. Clin. Exp. Pathol. 4 (5) (2011) 476–495.
[20] T T Zhang, Y Q Zhu, H Q Cai, et al., Prognostic role of aberrant mTOR activation in patients with stage II and III colorectal cancer, Biomark. Med. 14 (12) (2020) 1127–1137, https://doi.org/10.2217/bmm-2020-0141.
[21] S Jn, B Gl, E Al, Human DNAJ in cancer and stem cells, Cancer Lett. 312 (2) (2011), https://doi.org/10.1016/j.canlet.2011.08.019 [2022-11-24], https://pubmed.ncbi.nlm.nih.gov/21925790/.
[22] T T Zhang, Y Y Jiang, L Shang, et al., Overexpression of DNAJB6 promotes colorectal cancer cell invasion through an IQGAP1/ERK-dependent signaling pathway, Mol. Carcinog. 54 (10) (2015) 1205–1213, https://doi.org/10.1002/mc.22194.
[23] T Nagase, R Kikuno, K Ishikawa, et al., Prediction of the coding sequences of unidentified human genes. XVII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro, DNA Res. 7 (2) (2000) 143–150, https://doi.org/10.1093/dnares/7.2.143.
[24] X Yi, C Hu, C Zhang, et al., KIAA1522 is a new biomarker of promoting the tumorigenesis and distant metastasis of colorectal carcinoma, Cell. Signalling 90 (2022), 110202, https://doi.org/10.1016/j.cellsig.2021.110202.
[25] Y Z Liu, H Yang, J Cao, et al., KIAA1522 is a novel prognostic biomarker in patients with non-small cell lung cancer, Sci. Rep. 6 (2016) 24786, https://doi.org/10.1038/srep24786.
[26] Y Xu, C Sun, B Han, et al., High KIAA1522 expression predicts a poor prognosis in patients with hepatocellular carcinoma, Oncol. Lett. 20 (1) (2020) 509–516, https://doi.org/10.3892/ol.2020.11588.
[27] R L Camp, M Dolled-Filhart, D.L. Rimm, X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization, Clin. Cancer Res. 10 (21) (2004) 7252–7259, https://doi.org/10.1158/1078-0432.CCR-04-0713.
[28] P Mj, D. Rb, Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation, Stat. Med. 23 (13) (2004), https://doi.org/10.1002/sim.1802 [2022-11-28], https://pubmed.ncbi.nlm.nih.gov/15211606/.
[29] F E Harrell, K L Lee, D.B. Mark, Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors, Stat. Med. 15 (4) (1996) 361–387, https://doi.org/10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4.
[30] E Longato, M Vettoretti, B. Di Camillo, A practical perspective on the concordance index for the evaluation and selection of prognostic time-to-event models, J. Biomed. Inform. 108 (2020), 103496, https://doi.org/10.1016/j.jbi.2020.103496.
[31] J N Mandrekar, Receiver operating characteristic curve in diagnostic test assessment, J. Thorac. Oncol. 5 (9) (2010) 1315–1316, https://doi.org/10.1097/JTO.0b013e3181ec173d.
[32] J Wu, H Zhang, L Li, et al., A nomogram for predicting overall survival in patients with low-grade endometrial stromal sarcoma: a population-based analysis, Cancer Commun. 40 (7) (2020) 301–312, https://doi.org/10.1002/cac2.12067.
[33] G C Deng, Y Lv, H Yan, et al., Nomogram to predict survival of patients with advanced and metastatic pancreatic Cancer, BMC Cancer 21 (1) (2021) 1227, https://doi.org/10.1186/s12885-021-08943-w.
[34] A J Vickers, B van Calster, E.W. Steyerberg, A simple, step-by-step guide to interpreting decision curve analysis, Diagn. Progn. Res. 3 (2019) 18, https://doi.org/10.1186/s41512-019-0064-7.
[35] B Van Calster, L Wynants, J F M Verbeek, et al., Reporting and interpreting decision curve analysis: a guide for investigators, Eur. Urol. 74 (6) (2018) 796–804, https://doi.org/10.1016/j.eururo.2018.08.038.
[36] L Chen, D Lu, K Sun, et al., Identification of biomarkers associated with diagnosis and prognosis of colorectal cancer patients based on integrated bioinformatics analysis, Gene 692 (2019) 119–125, https://doi.org/10.1016/j.gene.2019.01.001.
[37] J Zhang, X Miao, T Wu, et al., Development and validation of ten-RNA binding protein signature predicts overall survival in osteosarcoma, Front. Mol. Biosci. 8 (2021), 751842, https://doi.org/10.3389/fmolb.2021.751842.
[38] P Ahluwalia, A K Mondal, C Bloomer, et al., Identification and clinical validation of a novel 4 gene-signature with prognostic utility in colorectal Cancer, Int. J. Mol. Sci. 20 (15) (2019) 3818, https://doi.org/10.3390/ijms20153818.
[39] B Zhu, P Zhang, M Liu, et al., Prognostic Significance of CSN2, CD8, and MMR status-associated nomograms in patients with colorectal cancer, Transl. Oncol. 11 (5) (2018) 1202–1212, https://doi.org/10.1016/j.tranon.2018.07.006.
[40] J Qu, Y Jiang, H Liu, et al., Prognostic value of E-cadherin-, CD44-, and MSH2-associated nomograms in patients with stage II and III colorectal cancer, Transl. Oncol. 10 (2) (2017) 121–131, https://doi.org/10.1016/j.tranon.2016.12.005.
[41] T Konishi, H Yoshidome, T Shida, et al., Phosphorylated mTOR expression as a predictor of survival after liver resection for colorectal liver metastases, J. Surg. Oncol. 124 (4) (2021) 598–606, https://doi.org/10.1002/jso.26551.
[42] D N Zouki, I Giannopoulou, P T Alexandrou, et al., MTOR/4EBP1 signaling and MMR status in colorectal cancer: new correlations and arising perspectives, Pathol. Res. Pract. 228 (2021), 153655, https://doi.org/10.1016/j.prp.2021.153655.

[43] T Sticz, A Molnár, Á Márk, et al., mTOR activity and its prognostic significance in human colorectal carcinoma depending on C1 and C2 complex-related protein expression, J. Clin. Pathol. 70 (5) (2017) 410–416, https://doi.org/10.1136/jclinpath-2016-203913.
[44] N Melling, R Simon, J R Izbicki, et al., Expression of phospho-mTOR kinase is abundant in colorectal cancer and associated with left-sided tumor localization, Int. J. Clin. Exp. Pathol. 8 (6) (2015) 7009–7015.
[45] Y J Zhang, X Q Tian, D F Sun, et al., Combined inhibition of MEK and mTOR signaling inhibits initiation and progression of colorectal cancer, Cancer Invest. 27 (3) (2009) 273–285, https://doi.org/10.1080/07357900802314893.
[46] P Gulhati, K A Bowen, J Liu, et al., mTORC1 and mTORC2 regulate EMT, motility, and metastasis of colorectal cancer via RhoA and Rac1 signaling pathways, Cancer Res. 71 (9) (2011) 3246–3256, https://doi.org/10.1158/0008-5472.CAN-10-4058.
[47] J Wang, D Liang, X P Zhang, et al., Novel PI3K/Akt/mTOR signaling inhibitor, W922, prevents colorectal cancer growth via the regulation of autophagy, Int. J. Oncol. 58 (1) (2020) 70–82, https://doi.org/10.3892/ijo.2020.5151.
[48] S Jiang, Y Zhang, Q Li, et al., KIAA1522 promotes the progression of hepatocellular carcinoma via the activation of the Wnt/β-Catenin signaling pathway, Onco Targets Ther. 13 (2020) 5657–5668, https://doi.org/10.2147/OTT.S251157.
[49] B Zhang, J Wang, X Wang, et al., Proteogenomic characterization of human colon and rectal cancer, Nature 513 (7518) (2014) 382–387, https://doi.org/10.1038/nature13438.
[50] J H Kim, K J Lee, Y Seo, et al., Effects of metformin on colorectal cancer stem cells depend on alterations in glutamine metabolism, Sci. Rep. 8 (1) (2018) 409, https://doi.org/10.1038/s41598-017-18762-4.