Genome-Wide Identification of a Novel Autophagy-Related Signature for Colorectal Cancer

Zhi Huang¹, Jie Liu¹, Liang Luo¹, Pan Sheng¹, Biao Wang¹, Jun Zhang¹, and Sha-sha Peng²,³

Abstract

Background: Plenty of evidence has suggested that autophagy plays a crucial role in the biological processes of cancers. This study aimed to screen autophagy-related genes (ARGs) and establish a novel a scoring system for colorectal cancer (CRC).

Methods: Autophagy-related genes sequencing data and the corresponding clinical data of CRC in The Cancer Genome Atlas were used as training data set. The GSE39582 data set from the Gene Expression Omnibus was used as validation set. An autophagy-related signature was developed in training set using univariate Cox analysis followed by stepwise multivariate Cox analysis and assessed in the validation set. Then we analyzed the function and pathways of ARGs using Gene Ontology and Kyoto Encyclopedia of Genes and Genomes (KEGG) database. Finally, a prognostic nomogram combining the autophagy-related risk score and clinicopathological characteristics was developed according to multivariate Cox analysis.

Results: After univariate and multivariate analysis, 3 ARGs were used to construct autophagy-related signature. The KEGG pathway analyses showed several significantly enriched oncological signatures, such as p53 signaling pathway, apoptosis, human cytomegalovirus infection, platinum drug resistance, necroptosis, and ErbB signaling pathway. Patients were divided into high- and low-risk groups, and patients with high risk had significantly shorter overall survival (OS) than low-risk patients in both training set and validation set. Furthermore, the nomogram for predicting 3- and 5-year OS was established based on autophagy-based risk score and clinicopathologic factors. The area under the curve and calibration curves indicated that the nomogram showed well accuracy of prediction.

Conclusions: Our proposed autophagy-based signature has important prognostic value and may provide a promising tool for the development of personalized therapy.

Keywords
colorectal cancer, autophagy-related genes, prognostic signature, RNA sequencing

¹ Department of General Surgery, Dazhou Central Hospital, Sichuan, People’s Republic of China
² Department of Hepatobiliary and Pancreatic Surgery, Huangshi Central Hospital of Edong Healthcare Group, Hubei Polytechnic University, Hubei, People’s Republic of China
³ Hubei Province Key Laboratory of Occupational Hazard Identification and Control, Wuhan University of Science and Technology, Wuhan, Hubei, People’s Republic of China

Received 12 October 2019; received revised 12 November 2019; accepted 19 November 2019

Corresponding Author:
Sha-sha Peng, Department of Hepatobiliary and Pancreatic Surgery, Huangshi Central Hospital of Edong Healthcare Group, Hubei Polytechnic University; Hubei Province Key Laboratory of Occupational Hazard Identification and Control, Wuhan University of Science and Technology, No. 141, Tianjin Road, Huangshi, Hubei 435000, People’s Republic of China.
Email: drpengss@126.com

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (http://www.creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
**Introduction**

Colorectal cancer (CRC) is one of the most frequently diagnosed malignancies worldwide. According to the 2018 Global Cancer Statistics, CRC is the third most common cancer worldwide and leading cause of cancer-related deaths.\(^1\) Despite recent advances in curative surgery and adjuvant therapy, the mortality is still high due to its late detection and high recurrence rates, and the 5-year survival rate of patients remains poor.\(^2\)-\(^4\) Therefore, it is urgently necessary to develop sensitive, cost-effective, and noninvasive methods to complement and improve current screening strategies for CRC diagnosis and prognosis.

Autophagy is a cell self-digestive process, which is initiated from the formation of autophagosome to deliver the protein and cell organelle to lysosome for degradation. This catabolic process is involved in a variety of autophagy-related genes (ARGs), which could influence the activation of autophagy.\(^5\) Under some pathological conditions such as inflammation, neurodegeneration, aging, and tumor, autophagy can be stimulated to maintain cellular homeostasis.\(^6\),\(^7\) However, the role of autophagy in tumorigenesis remains controversial, and mechanism is still rudimentary and inconclusive. Since autophagy has complex functions in cancer, it is urgent to further explore the potential biological processes of autophagy in tumorigenesis and development.

The most commonly used approach for predicting patient survival remains pathological staging according to the tumor–node–metastasis (TNM) classification system.\(^8\) However, this staging system is not sufficient to assess prognosis and does not reflect the biological heterogeneity of cancer. It provides only limited information for the clinical prognostic since even patients within the same stage display a strong heterogeneity for prognosis and treatment response. Other factors such as age, performance status, and tumor location can also influence the patients’ survival.\(^9\)-\(^11\) In addition, this system is not very useful in predicting individual patient outcomes. Therefore, this highlights the critical need to develop robust prognostic biomarkers that can offer a superior prognostic clinical usefulness. In recent years, nomograms have gradually gained popularity. They are valuable tools to use different clinical variables to determine a statistical prognostic model that generates a probability of clinical outcomes for a single patient.\(^12\) Nomograms have been applied in various types of cancers.\(^13\),\(^14\) For example, Zhao et al.\(^14\) used 5 bone metastases-related genes to establish a prognostic nomogram, which can help in prediction of early bone metastases for patients with breast cancer. Similarly, Peak et al.\(^15\) developed a nomogram to predict lymph node metastasis based on 3 variables, including tumor grade, lymphovascular invasion, and clinical lymph node status. To our best knowledge, no prior prognostic models for CRC were established based on a combination of ARGs and clinicopathological features. Therefore, the aim of the present study was to identify CRC-related autophagy genes and build a novel prognostic signature through a comprehensive analysis of TCGA sequencing data.

**Materials and Methods**

**Data Acquisition**

The autophagy gene list was downloaded from the Human Autophagy Database (HADb, http://autophagy.lu/clustering/index.html). The data of messenger RNA (mRNA)-sequencing and the corresponding clinical data in patients with CRC were collected from The Cancer Genome Atlas (TCGA) database. To validate the predictive ability of autophagy-related risk score, we obtained the GSE39582 data set from the Gene Expression Omnibus (GEO) database. Overall, 499 samples from the TCGA database were used as training set, 556 samples from GSE39582 data set were used as validation set. Approval by ethics committee would not be necessary because all data had been collected from public availability of data in the HADb, GEO, and TCGA databases.

**The Differentially Expressed ARGs Screening**

All genes were downloaded in FPKM format and the genes that met the cutoff criteria of \(P < .05\) and fold change (FC) > 2.0 were screened out as differentially expressed ARGs. The R package limma was used to identify differentially expressed ARGs between normal samples and tumor samples.

**Functional Enrichment Analysis**

The screening of the functions of differentially expressed ARGs was analyzed by the Gene Ontology (GO), which could obtain the biological function of gene. In addition, the Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis was used to enrich the differential pathway by “clusterProfiler” package in R/Bioconductor.\(^16\) The threshold for enrichment significance was \(P\) value \(< .05\).

**Signature Development and Validation**

A univariate Cox proportional hazards model was used to screen out ARGs that are significantly correlated with the overall survival (OS) of patients with CRC. These ARGs with a \(P\) value \(< .05\) by univariate analysis were subjected to a multivariate Cox regression analysis with stepwise selection of variables based on the Akaike information criterion for identifying optimal prognostic ARGs. A autophagy-related signature was constructed based on a linear combination of the relative expression level of genes multiplied by regression coefficients \((b)\), which represented the relative weight of genes in the multiple Cox regression analysis. The risk score \(= (\beta\text{mRNA}_1 \times \text{expression level of mRNA}_1) + (\beta\text{mRNA}_2 \times \text{expression level of mRNA}_2) + (\beta\text{mRNA}_3 \times \text{expression level of mRNA}_3) + \ldots + (\beta\text{mRNA}_n \times \text{expression level of mRNA}_n).\) Patients were divided into high-risk group and low-risk group by the median risk score as the cutoff value. Kaplan-Meier curves were plotted and survival differences between high-risk and low-risk group were compared by the log-rank test. The concordance index (C-index) was used to assess the prediction
accuracy of the risk score model. In addition, GSE14520 data set was downloaded from the GEO database as a validation set. The risk score for each patient was calculated with the same formula as training set. The Kaplan-Meier curve was used to assess the predictive ability of the autophagy-related signature.

**Construction and Validation of Nomogram**

To explore whether the autophagy-related signature could be independent of other clinicopathological parameters (including age, gender, tumor location, prior malignancy, pretreatment carcinoembryonic antigen (CEA), Kirsten rat sarcoma (KRAS) mutation, lymphatic invasion, tumor stage, primary therapy outcome, and radiation therapy), Cox proportional hazard regression was performed. Variables in the univariate analysis with \( P \) values <.05 were selected for multivariate analysis. A nomogram for predicting the 3-year and 5-year OS was established based on the result of multivariate Cox regression analysis. The internal validation of nomogram was performed by discrimination and calibration. The discrimination was estimated by the area under the curve (AUC) of receiver operating characteristic (ROC) curve.\(^{17}\) The AUC ranges from 0.5 to 1.0, with 0.5 indicates the outcomes is totally random and 1.0 indicates the perfect discrimination. The calibration was assessed by calibration curves, which shows how close the nomogram estimated risk was to the observed risk.

**Statistical Analyses**

Quantitative variables were compared using the \( t \) test or Mann-Whitney \( U \) test; the Pearson \( \chi^2 \) test or Fisher exact test was used to analyze categorical variables. Survival curves were created by the Kaplan-Meier method and compared using the log-rank test. The Cox regression coefficients were used to establish a risk score signature. The nomogram was built on the basis of multivariate analysis results by the package of “rms” in R. The prediction ability of nomogram was assessed by area under the ROC curve in the package of “survivalROC” in R. Calibration plots were used to compare the observed and predicted probabilities for the nomogram. All the statistical analyses were performed by SPSS 23.0 (Chicago, Illinois),
along with version 3.5.3 of R software. Data were considered to be statistically significant when \( P \) value was less than .05.

**Results**

**Identification of Differentially Expressed ARGs**

The overall process of this study is described as Figure 1. In this work, we collected an mRNA expression data from TCGA data set, and it contained 571 CRC samples and 44 normal samples. A total of 231 ARGs were obtained from the HADb. Based on the cutoff criteria (\( P \) value < .05 and FC > 2.0), we identified 36 differentially expressed ARGs, including 16 upregulated and 20 downregulated genes. The expression of 36 differentially expressed ARGs between CRC and normal tissues was visualized by scatter plots (Figure 2). The clinicopathologic characteristics of patients with CRC are summarized in Table 1.

**Functional Enrichment Analysis**

Thirty-six differentially expressed ARGs were annotated according to the GO database. We enriched the functions of these mRNA (\( P < .05 \)). As shown in Figure 3A, autophagy and process utilizing autophagic mechanism were significantly enriched for biological processes. While for cellular component was autophagosome, for molecular functions were ubiquitin protein ligase binding and ubiquitin-like protein ligase binding. Furthermore, we analyzed KEGG pathways of these ARGs. We found certain KEGG pathways positively associated with ARGs such as p53 signaling pathway, apoptosis, human cytomegalovirus infection (HCMV), platinum drug resistance, necroptosis, and ErbB signaling pathway (Figure 3B).

**Establishment of Autophagy-Related Signature in Training Set**

To identify the OS-related ARGs, 36 differentially expressed ARGs were initially subjected to univariate Cox proportional hazards regression analysis in the training set. The result showed that 4 ARGs were significantly associated with the OS (\( P < .05 \)). Then the 4 prognosis-related ARGs were subsequently selected for multivariate Cox proportional hazards regression. Finally, 3 ARGs (CDKN2A, NRG1, and VEGFA) were confirmed and used to establish an autophagy-related signature (Table 2). These include 2 potential risky genes and 1 potential protective gene. The risk score for OS was calculated as follows: risk score = \( \frac{0.203082}{C_{CDKN2A}} + \frac{1.21191}{C_{NRG1}} + \frac{0.321454}{C_{VEGFA}} \). We then calculated the risk score for each patient used this signature and used the median risk score as a cutoff value for classifying patient into a high-risk group (n = 249) and a low-risk group (n = 250), respectively. The heatmap, the risk score, and OS time of these 3 prognostic ARGs in training set are shown in Figure 4A. Kaplan-Meier curve and time-dependent ROC curves were used to evaluate the prognostic potential of the 3 autophagy-related signature. Kaplan-Meier curves showed that patients with high-risk group had a significantly lower OS than patients with the low-risk group (\( P < .001 \); Figure 4A). The C-index for OS was 0.784 for training set.

**Validation of the Autophagy-Related Signature**

To validate the predictive ability of autophagy-related signature, risk scores were calculated with the same formula for patient in GSE39582. Consistent with the results in the TCGA set, patients were categorized into high-risk group and low-risk...
Table 1. Clinicopathologic Characteristics of Patients With CRC in TCGA and GEO Set.*

| Variables                      | TCGA (n = 499) | GEO (n = 556) |
|-------------------------------|---------------|--------------|
| Age (mean ± SD, years)        | 66.12 ± 12.62 | 66.73 ± 13.29 |
| Gender                        |               |              |
| Female                        | 229 (45.9)    | 249 (44.8)   |
| Male                          | 270 (54.1)    | 307 (55.2)   |
| Primary tumor location        |               |              |
| Colon                         | 315 (63.1)    | 556 (100)    |
| Rectum                        | 80 (16.0)     | /            |
| Unknown                       | 104 (20.8)    | /            |
| Prior malignancy              |               |              |
| No                            | 436 (87.4)    | /            |
| Yes                           | 63 (12.6)     | /            |
| Pretreatment CEA level (ng/mL)| 70.96 ± 507.53| /            |
| KRAS mutation                 |               |              |
| Wild type                     | 26 (5.2)      | 322 (57.9)   |
| Mutant                        | 24 (4.8)      | 213 (38.3)   |
| Unknown                       | 449 (90.0)    | 21 (3.8)     |
| Lymphatic invasion            |               |              |
| No                            | 270 (54.1)    | /            |
| Yes                           | 51 (10.2)     | /            |
| Unknown                       | 178 (35.7)    | /            |
| Stage                         |               |              |
| I                             | 90 (18.0)     | 36 (6.5)     |
| II                            | 193 (38.7)    | 258 (46.4)   |
| III                           | 142 (28.5)    | 203 (36.5)   |
| IV                            | 74 (14.8)     | 59 (10.6)    |
| Primary therapy outcome       |               |              |
| Complete response             | 140 (28.1)    | /            |
| Other                         | 43 (8.6)      | /            |
| Unknown                       | 316 (63.3)    | /            |
| Radiotherapy                  |               |              |
| Yes                           | 23 (4.6)      | /            |
| No                            | 405 (81.2)    | /            |
| Unknown                       | 71 (14.2)     | /            |

Abbreviations: CEA, carcinoembryonic antigen; CRC, colorectal cancer; GEO, Gene Expression Omnibus; KRAS, Kirsten rat sarcoma; SD, standard deviation; TCGA, The Cancer Genome Atlas.

*Other: Partial response, progressive disease, and stable disease.

Discussion

Colorectal cancer remains a huge health burden worldwide with substantial mortality. The TNM staging system is routinely used as a staging procedure for patients with CRC. However, significant heterogeneity for prognosis and treatment response still exists in the same stage. An accurate prediction of prognosis may inform clinical decisions. In recent years, high-throughput biotechnology is widely used to identify genes as markers for CRC. It suggested that a deeper understanding of the genetic characteristics of CRC may lead to a better risk stratification system. Furthermore, many studies have explored the role of autophagy mechanisms in tumorigenesis, including CRC. However, most studies focus on autophagy only by studying a single gene. To our best knowledge, no autophagy-based signature for CRC was constructed. Therefore, we aimed to identify ARGs associated with OS of patients with CRC and establish a novel autophagy-related prognostic model through the TCGA and GEO databases.

In present study, we identified 36 ARGs, which were dysregulated in patients with CRC. Considering these genes may be involved in the initiation of CRC, we further analyzed these different expression ARG functions through the GO and KEGG pathway analysis. The result demonstrated several significantly enriched oncological signatures, such as p53 signaling pathway, apoptosis, HCMV, platinum drug resistance, necroptosis, and ErbB signaling pathway, which might help explain the underlying molecular mechanisms of CRC. After univariate and multivariate analysis, 3 ARGs were used to construct risk score signature using their expression levels weighted by corresponding correlation coefficient. The autophagy-related signature could divide patients into high-risk and low-risk groups based on the median risk score. Patients with high-risk score have significantly worse OS than patients in low-risk group. Furthermore, the predictive accuracy of autophagy-related signature...
was successfully validated in the cohorts of GSE39582 data set, indicating the good reproducibility of this signature. Multivariate analyses demonstrated that the risk score was independent prognostic factor for patients with CRC. Finally, a prognostic nomogram was established based on risk score and 3 clinicopathologic factors. The prognostic accuracy of the nomogram was confirmed by the discrimination and calibration plots.

Our study has several merits. First, to our knowledge, we firstly established an autophagy-based signature based on multi–data sets. The signature showed favorable predictive ability. Second, the results of the TCGA training data set (mainly composed of the US and European populations) are well validated in the validation set of the Asian population composition. Validation in multiple countries enhances the reliability of the results, indicating that ATG scores may be appropriate for patients of different ethnicities. Third, our nomogram combining risk score with conventional clinical parameters shown significantly improved performance. For example, the nomogram showed a higher predictive ability than the AJCC TNM staging system, consistent with previous studies. Moreover, nomogram is of great significance to clinicians and patients because that it can predict individual patient outcomes. When a patient who was diagnosed with a tumor for a few several years previously comes to the clinician for help, he is more concerned about his individual survival risk since at this particular moment rather than the traditional risk since diagnosis.

The proposed autophagy-related signature included 3 ARGs (VEGFA, NRG1, and CDKN2A). All genes in the signature were associated with OS of patients with CRC. Of these, VEGFA, NRG1, and CDKN2A have been previously reported to have a clear correlation with CRC. Vascular epithelial growth factor A (VEGFA), as a member of the VEGF family, is involved in tumor growth, metastasis, and angiogenesis. Several studies indicated that VEGFA is a direct target gene for many microRNAs and was closely related to the initiation and progression of CRC. For example, Chen et al demonstrated that miR-150-5p inhibits the progression of CRC by targeting VEGFA in CRC as a tumor suppressor. In addition, the level of VEGF is associated with clinical outcomes. Therefore, VEGF has been the main focus of anticancer drug development. Studies have shown that antiangiogenic therapy can effectively improve the survival rate of patients with CRC, which indicates that inhibition of tumor angiogenesis is a potential method for the treatment of CRC. NRG1 is a membrane glycoprotein that mediates cellular signaling and plays a key role in the growth and development of many human organs. This gene can be used as an oncogene or as a tumor suppressor gene to participate in the development of CRC. The gene produces a number of different isoforms by replacing the use of promoters and splicing. These isoforms are expressed in a tissue-specific manner with significant structural differences and are classified into at least 3 subgroups (types I-III). This gene can be used as an oncogene or as a tumor suppressor gene to participate in the development of CRC. De Boeck et al found that the expression of NRG1 was significantly upregulated in the CRC tissues and was linked to advanced stage and invasion depth.

### Table 2. Three Prognostic ARGs Significantly Associated With OS in the TCGA Set.

| Gene Name | Coefficient | Downregulated/Upregulated | Hazard Ratio (95% Confidence Interval) | P Value |
|-----------|-------------|----------------------------|----------------------------------------|---------|
| CDKN2A    | 0.203082    | Up                         | 1.23 (1.04-1.45)                       | .017    |
| NRG1      | -1.21191    | Down                       | 0.29 (0.10-0.85)                       | .023    |
| VEGFA     | 0.321454    | Up                         | 1.38 (1.02-1.87)                       | .037    |

Abbreviations: ARG, autophagy related gene; OS, overall survival; TCGA, The Cancer Genome Atlas.
NRG1 promotes CRC tumorigenesis by activating the HER2/HER3-dependent PI3K/AKT signaling cascade in CRC cells. On the contrary, there are reports that NRG1 may be the major tumor suppressor gene that causes 8p loss in the colon (short arm of chromosome 8). In our study, we found that NRG1 was downregulated in CRC and associated with worse OS. Therefore, the exact relationship between the NRG1 and CRC needs to be further investigated in future experiments.

CDKN2A is one of the original CIMP marker group genes and has important functions in carcinogenesis. Several studies indicated that CDKN2A hypermethylation might be a predictive factor for patients with CRC.

In our study, functional enrichment analysis indicated that ARGs were identified in several CRC-related pathways including the following: p53 signaling pathway, apoptosis, HCMV, necroptosis, and ErbB signaling pathway. P53 is a classic tumor suppressor that often mutates between multiple cancer types. As a key guard in tumorigenesis, p53 is activated in response to carcinogenesis and regulates a large number of genes.

**Figure 4.** The distribution of risk score, survival status, the heatmap, and Kaplan-Meier survival curves of 3-autophagy-related signature in training set (A) and in the validation set (B).
Table 3. Univariate and Multivariate Analyses of Overall Survival in TCGA Set.

| Variable                        | Univariate Analysis       |          | Multivariate Analysis      |          |
|---------------------------------|---------------------------|----------|---------------------------|----------|
|                                 | HR (95% CI)               | P Value  | HR (95% CI)               | P Value  |
| Age                             | 1.031 (1.012-1.050)       | <.001    | 1.051 (1.023-1.079)       | <.001    |
| Gender                          |                           |          |                           |          |
| Female                          | 1.041 (0.699-1.551)       | .843     |                           |          |
| Male                            |                           |          |                           |          |
| Primary tumor location          | 0.905 (0.690-1.186)       | .468     |                           |          |
| Colon                           |                           |          |                           |          |
| Rectum                          |                           |          |                           |          |
| Unknown                         |                           |          |                           |          |
| Prior malignancy                | 1.309 (0.765-2.238)       | .325     |                           |          |
| No                              |                           |          |                           |          |
| Yes                             |                           |          |                           |          |
| Pretreatment CEA level (ng/mL)  | 1.000 (1.000-1.000)       | .020     | 1.000 (1.000-1.001)       | .337     |
| KRAS mutation                   |                           |          |                           |          |
| Wild-type                       | 0.716 (0.536-0.958)       | .024     |                           |          |
| Mutant                          | 1.877 (0.347-10.152)      | .465     |                           |          |
| Unknown                         | 0.974 (0.271-3.498)       | .968     |                           |          |
| Lymphatic invasion              | 1.420 (1.086-1.857)       | .010     |                           |          |
| No                              |                           |          |                           |          |
| Yes                             | 0.944 (0.476-1.872)       | .870     |                           |          |
| Unknown                         | 0.778 (0.202-3.001)       | .716     |                           |          |
| Stage                           | 2.170 (1.725-2.730)       | <.001    |                           |          |
| I                               |                           |          |                           |          |
| II                              | 1.285 (0.395-4.182)       | .677     |                           |          |
| III                             | 4.237 (1.196-15.007)      | .025     |                           |          |
| IV                              | 7.953 (2.332-27.119)      | <.001    |                           |          |
| Primary therapy outcome         | 1.415 (1.091-1.835)       | .009     |                           |          |
| Complete response               |                           |          |                           |          |
| Other                           | 2.908 (1.076-7.863)       | .035     |                           |          |
| Unknown                         | 0.967 (0.415-2.255)       | .939     |                           |          |
| Radiotherapy                    | 2.082 (1.352-3.206)       | <.001    |                           |          |
| Yes                             |                           |          |                           |          |
| No                              | 4.545 (0.518-39.863)      | .172     |                           |          |
| Unknown                         | 49.227 (4.242-571.231)    | .002     |                           |          |

Abbreviations: CI, confidence interval; HR, hazard ratio; TCGA, The Cancer Genome Atlas.

Figure 5. Nomogram for predicting 3- and 5-year overall survival of patients with colorectal cancer.
involved in cell cycle arrest, DNA repair, and apoptosis. Mutations of p53 are highly frequent in CRC, and p53 mutation status is associated with disease outcome. A meta-analysis demonstrated that HCMV infection was associated with an increased risk of CRC. Teo et al found that HCMV infection can enhance cell proliferation, migration, and upregulation of epithelial–mesenchymal transition markers in CRC cells. Taken together, these differentially expressed ARGs are involved in many important CRC-associated biological functions and pathways.

There are still some limitations of this study. First, the sample size is relatively small, and some variables that may be important do not show independent prognostic value. Therefore, future studies with large scale may build more accurate models. Second, due to the lack of primary therapy outcome variable for GSE39582 data set, we are unable to evaluate the predictive accuracy of nomogram by external validation. Third, all analyses in our study are descriptive, and further functional experiments are needed to explore the molecular mechanisms of these 3 ARGs.

Conclusions
We conducted a novel 3 autophagy-related signature to predict OS of patients with CRC, which may help in clinical decision-making for individualized therapy.

Authors’ Note
Z.H. and J.L. contributed equally to this work. Z.H. and S.S.P. designed the experiment. Z.H., J.L., L.L., and P.S. undertook the data acquisition. Z.H., J.L., P.S., J.L., and B.W. were involved in the interpretation of data. Z.H., J.L., L.L., J.Z., and S.S.P. analyzed and visualized the data. All authors drafted and revised the manuscript. The final manuscript was read and approved by all authors. The data sets

Figure 6. Comparison of the areas under the curve (AUCs) of the nomogram and American Joint Committee on Cancer (AJCC) tumor–node–metastasis (TNM) staging system. A: 3-year OS; B: 5-year OS.

Figure 7. Calibration curves of the nomogram prediction of (A) 3-year and (B) 5-year survival of patients with colorectal cancer.
used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Health commission of Hubei Province Scientific Research Project (WJ2019H160).

ORCID iD
Sha-sha Peng https://orcid.org/0000-0001-7808-7263

References
1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
2. Ludmir EB, Palta M, Willett CG, Czito BG. Total neoadjuvant therapy for rectal cancer: an emerging option. Cancer. 2017;123(9):1497-1506.
3. Roxburgh CS, McMillan DC. The role of the in situ local inflammatory response in predicting recurrence and survival in patients with primary operable colorectal cancer. Cancer Treat Rev. 2012;38(5):451-466.
4. Van Cutsem E, Nordlinger B, Adam R, et al. Towards a pan-European consensus on the treatment of patients with colorectal liver metastases. Eur J Can (Oxford, England: 1990). 2006;42(14):2212-2221.
5. Weidberg H, Shvets E, Elazar Z. Biogenesis and cargo selectivity of autophagosomes. Annu Rev Biochem. 2011;80:125-156.
6. Levine B, Kroemer G. Autophagy in the pathogenesis of disease. Cell. 2008;132(1):27-42.
7. Mizushima N, Levine B, Cuervo AM, Klionsky DJ. Autophagy fights disease through cellular self-digestion. Nature. 2008;451(7182):1069-1075.
8. Amin MB, Greene FL, Edge SB, et al. The Eighth Edition AJCC Cancer Staging Manual: continuing to build a bridge from a population-based to a more “personalized” approach to cancer staging. CA Cancer J Clin. 2017;67(2):93-99.
9. Fu J, Ruan H, Zheng H, et al. Impact of old age on resectable colorectal cancer outcomes. PeerJ. 2019;7:e6350.
10. Jehn CF, Bönig L, Krönig H, Pezzutto A, Lüftner D. Influence of comorbidity, age and performance status on treatment efficacy and safety of cetuximab plus irinotecan in irinotecan-refractory elderly patients with metastatic colorectal cancer. Eur J Cancer. 2014;50(7):1269-1275.
11. Wang CB, Shahjehan F, Merchea A, et al. Impact of tumor location and variables associated with overall survival in patients with colorectal cancer: A Mayo Clinic Colon and Rectal Cancer Registry Study. Front Oncol. 2019;9:76.
12. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. Lancet Oncol. 2015;16(4):e173-e180.
13. Dong D, Tang L, Li ZY, et al. Development and validation of an individualized nomogram to identify occult peritoneal metastasis in patients with advanced gastric cancer. Ann Oncol. 2019;30(3):431-438.
14. Zhao C, Lou Y, Wang Y, et al. A gene expression signature-based nomogram model in prediction of breast cancer bone metastases. Cancer Med. 2019;8(1):200-208.
15. Peak TC, Russell GB, Dutta R, Rotheberg MB, Chapple AG, Hemal AK. A national cancer database-based nomogram to predict lymph node metastasis in penile cancer. BJU Int. 2019;123(6):1005-1010.
16. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284-287.
17. Wolbers M, Koller MT, Witteman JC, Steyerberg EW. Prognostic models with competing risks: methods and application to coronary risk prediction. Epidemiology. 2009;20(4):555-561.
18. Moody L, Dworetsky S, An R, Mantha S, Pan YX. The efficacy of miR-20a as a diagnostic and prognostic biomarker for colorectal cancer: a systematic review and meta-analysis. Cancers (Basel). 2019;11(8):E1111.
19. Pan H, Pan J, Song S, Ji L, Lv H, Yang Z. EXOSC5 as a novel prognostic marker promotes proliferation of colorectal cancer via activating the ERK and AKT pathways. Front Oncol. 2019;9:643.
20. Wang D, Yuan W, Wang Y, et al. Serum CCL20 combined with IL-17A as early diagnostic and prognostic biomarkers for human colorectal cancer. J Transl Med. 2019;17(1):253.
21. Fu Y, Hong L, Xu J, et al. Discovery of a small molecule targeting autophagy via ATG4B inhibition and cell death of colorectal cancer cells in vitro and in vivo. Autophagy. 2019;15(2):295-311.
22. Wang M, Han D, Yuan Z, et al. Long non-coding RNA H19 confers 5-Fu resistance in colorectal cancer by promoting SIRT1-mediated autophagy. Cell Death Dis. 2018;9(12):1149.
23. Zhou L, Gao W, Wang K, et al. Brefeldin A inhibits colorectal cancer growth by triggering Bip/Akt-regulated autophagy. FASEB J. 2019;33(4):5520-5534.
24. Li J, Gu J, Ma X, et al. Development and validation of a nomogram for predicting survival in Chinese Han patients with resected colorectal cancer. J Surg Oncol. 2018;118(6):1034-1041.
25. Yu C, Zhang Y. Development and validation of a prognostic nomogram for early-onset colon cancer. Biosci Rep. 2019;39(6):BSR20181781.
26. Chen X, Zeng K, Xu M, et al. SP1-induced lncRNA-ZFAS1 contributes to colorectal cancer progression via the miR-150-5p/VEGFA axis. Cell Death Dis. 2018;9(10):982.
27. Qiu Y, Yu H, Shi X, et al. microRNA-497 inhibits invasion and metastasis of colorectal cancer cells by targeting vascular endothelial growth factor-A. Cell Prolif. 2016;49(1):69-78.
28. Wang JY, Jiang JB, Li Y, Wang YL, Dai Y. MicroRNA-299-3p suppresses proliferation and invasion by targeting VEGFA in human colon carcinoma. Biomed Pharmacother. 2017;93:1047-1054.
29. Yang X, Qiu J, Kang H, Wang Y, Qian J. miR-125a-5p suppresses colorectal cancer progression by targeting VEGFA. *Cancer Manag Res*. 2018;10:5839-5853.

30. Chen X, Xu X, Pan B, et al. miR-150-5p suppresses tumor progression by targeting VEGFA in colorectal cancer. *Aging (Albany NY)*. 2018;10(11):3421-3437.

31. Takahashi N, Iwasa S, Taniguchi H, et al. Prognostic role of ERBB2, MET and VEGFA expression in metastatic colorectal cancer patients treated with anti-EGFR antibodies. *Br J Cancer*. 2016;114(9):1003-1011.

32. Mésange P, Bouygues A, Ferrand N, et al. Combinations of bevacizumab and erlotinib show activity in colorectal cancer independent of status. *Clin Cancer Res*. 2018;24(11):2548-2558.

33. Tournigand C, Chibaudel B, Samson B, et al. Bevacizumab with or without erlotinib as maintenance therapy in patients with metastatic colorectal cancer (GERCOR DREAM; OPTIMOX3): a randomised, open-label, phase 3 trial. *Lancet Oncol*. 2015;16(15):1493-1505.

34. Liu X, Bates R, Yin DM, et al. Specific regulation of NRG1 isoform expression by neuronal activity. *J Neurosci*. 2011;31(23):8491-8501.

35. Steinthorsdottir V, Stefansson H, Ghosh S, et al. Multiple novel transcription initiation sites for NRG1. *Gene*. 2004;342(1):97-105.

36. Tan W, Wang Y, Gold B, et al. Molecular cloning of a brain-specific, developmentally regulated neuregulin 1 (NRG1) isoform and identification of a functional promoter variant associated with schizophrenia. *J Biol Chem*. 2007;282(33):24343-24351.

37. Zhao WJ. The expression and localization of neuregulin-1 (Nrg1) in the gastrointestinal system of the rhesus monkey. *Folia Histochem Cytobiol*. 2013;51(1):38-44.

38. De Boeck A, Pauwels P, Hensen K, et al. Bone marrow-derived mesenchymal stem cells promote colorectal cancer progression through paracrine neuregulin 1/HER3 signalling. *Gut*. 2013;62(4):550-560.

39. Pole JC, Courty-Cahen C, Garcia MJ, et al. High-resolution analysis of chromosome rearrangements on 8p in breast, colon and pancreatic cancer reveals a complex pattern of loss, gain and translocation. *Oncogene*. 2006;25(41):5693-5706.

40. Toyota M, Issa JP. CpG island methylator phenotypes in aging and cancer. *Semin Cancer Biol*. 1999;9(5):349-357.

41. Kohonen-Corish MR, Tseung J, Chan C, et al. KRAS mutations and CDKN2A promoter methylation show an interactive adverse effect on survival and predict recurrence of rectal cancer. *Int J Cancer*. 2014;134(12):2820-2828.

42. Xing X, Cai W, Shi H, et al. The prognostic value of CDKN2A hypermethylation in colorectal cancer: a meta-analysis. *Br J Cancer*. 2013;108(12):2542-2548.

43. Russo A, Bazan V, Iacopetta B, Kerr D, Soussi T, Gebbia N. The TP53 colorectal cancer international collaborative study on the prognostic and predictive significance of p53 mutation: influence of tumor site, type of mutation, and adjuvant treatment. *J Clin Oncol*. 2005;23(30):7518-7528.

44. Wade M, Li YC, Wahl GM. MDM2, MDMX and p53 in oncogenesis and cancer therapy. *Nat Rev Cancer*. 2013;13(2):83-96.

45. Bai B, Wang X, Chen E, Zhu H. Human cytomegalovirus infection and colorectal cancer risk: a meta-analysis. *Oncotarget*. 2016;7(47):76735-76742.

46. Teo WH, Chen HP, Huang JC, Chan YJ. Human cytomegalovirus infection enhances cell proliferation, migration and upregulation of EMT markers in colorectal cancer-derived stem cell-like cells. *Int J Oncol*. 2017;51(5):1415-1426.