

RESEARCH

Open Access

Development of prognosis model for colon cancer based on autophagy-related genes



Xu Wang, Yuanmin Xu, Ting Li, Bo Chen and Wenqi Yang*

Abstract

Background: Autophagy is an orderly catabolic process for degrading and removing unnecessary or dysfunctional cellular components such as proteins and organelles. Although autophagy is known to play an important role in various types of cancer, the effects of autophagy-related genes (ARGs) on colon cancer have not been well studied.

Methods: Expression profiles from ARGs in 457 colon cancer patients were retrieved from the TCGA database (<https://portal.gdc.cancer.gov>). Differentially expressed ARGs and ARGs related to overall patient survival were identified. Cox proportional-hazard models were used to investigate the association between ARG expression profiles and patient prognosis.

Results: Twenty ARGs were significantly associated with the overall survival of colon cancer patients. Five of these ARGs had a mutation rate $\geq 3\%$. Patients were divided into high-risk and low-risk groups based on Cox regression analysis of 8 ARGs. Low-risk patients had a significantly longer survival time than high-risk patients ($p < 0.001$). Univariate and multivariate Cox regression analysis showed that the resulting risk score, which was associated with infiltration depth and metastasis, could be an independent predictor of patient survival. A nomogram was established to predict 1-, 3-, and 5-year survival of colon cancer patients based on 5 independent prognosis factors, including the risk score. The prognostic nomogram with online webserver was more effective and convenient to provide information for researchers and clinicians.

Conclusion: The 8 ARGs can be used to predict the prognosis of patients and provide information for their individualized treatment.

Keywords: Autophagy-related genes, Prognosis model, Colon cancer, TCGA

Introduction

Despite rapid advances in medical science and technology, cancer incidence and cancer-related mortality rates are increasing rapidly worldwide [1]. Patients with early stage colon cancer can be successfully treated by surgery; however, most patients with advanced colon cancer experience recurrence and metastasis and typically exhibit 5-year survival rates $< 10\%$ [2–4]. Although tumor size, stage, and histological grade are often used to predict prognosis of colon cancer patients, these indicators do

not accurately predict patient survival and are not useful for developing individualized treatment regimens. With the development of chemotherapy and targeted therapeutics, the overall survival rate of colon cancer patients has increased significantly. Carcinoembryonic antigen (CEA) has been widely used in colon cancer diagnosis, but more efficient molecular biomarkers for early diagnosis and advanced therapeutic agents are needed to improve prognosis and treatment outcomes in colon cancer patients [5].

Autophagy is a multi-step process of intracellular degradation closely controlled by numerous ARGs, which occurs under a variety of stress conditions, including organelle damage, the presence of abnormal proteins, and

* Correspondence: doctoryangwenqi@sina.com
Department of General Surgery, The First Affiliated Hospital of Anhui Medical University, Hefei 230032, Anhui, China



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

nutritional deficiency [6]. Autophagy plays an important role in various aspects of tumor suppression, including the response of cells to nutrition and hypoxia stress, control of programmed cell death, and tumor-related immune response. Under normal physiological conditions, autophagy keeps cells in a stable state, prevents the accumulation of damaged and potentially carcinogenic proteins and organelles, and inhibits the process of carcinogenesis. However, once tumors begin to form, autophagy provides an abundance of nutrients for cancer cells and promotes tumor growth [7].

Over the past two decades, abundant researches have provided important information on the correlation between autophagy and colon cancer [8–10]. For example, Schroll et al. [9] suggested that cancer cells may become more sensitive to chemotherapy in an environment of glucose restriction and autophagy inhibition. Autophagy is now widely recognized to play an important role in colon cancer growth and progression and may be useful in anti-cancer therapies [11]. Previous studies have focused primarily on relationships between one or several ARGs and colon cancer, but limited research has been devoted to large scale searches for ARGs related to patient prognosis. In this study, we contribute to this growing area of research by exploring the value of ARGs in predicting prognosis of colon cancer patients and improving clinical decision-making for individualized treatment. We used clinical data and large-scale patterns of ARG expression in colon cancer patients to develop an informative model of prognosis.

Materials and methods

ARG identification and expression

A list of 232 human ARGs was constructed using the Human Autophagy Database (HADb; <http://www.autophagy.lu>), a publicly available repository containing up-to-date information on human genes and proteins that are directly or indirectly involved in autophagy. Expression patterns for the 232 ARGs and clinical information from 457 colon cancer patients were downloaded from the Genomic Data Commons (GDC) Data Portal (<https://portal.gdc.cancer.gov>).

GO and KEGG analysis

To better understand the biological functions of the ARGs, gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) analyses were performed using “ggplot2,” “Bioconductor,” and “org.Hs.eg.db” R packages. GO and KEGG pathways with p values and q values < 0.05 were considered to be significant.

Establishment of an ARG-related prognosis model

Twenty ARGs that were significantly related to patient prognosis ($p < 0.05$) were identified using univariate Cox regression analysis. The cBioPortal for Cancer Genomics

online website (<https://www.cbioportal.org>) was used to determine the mutation rate for each of the 20 ARGs. A risk score for each patient was calculated based on the expression of these ARGs using multivariate Cox regression analysis. Patients were then divided into high-risk and low-risk groups based on the risk score; the median risk score was used to differentiate the two groups. The Kaplan-Meier method was used to evaluate survival differences between the high- and low-risk groups, and the log-rank statistical method was used for comparison. Univariate and multivariate analyses were used to determine if the risk score was an independent predictor of prognosis in colon cancer patients. Receiver operator characteristic curves (ROC) and area under the curves (AUC) were used to test the prediction efficiency of the prognosis model. A nomogram was established based on five independent prognosis factors that were significant in both the univariate and multivariate analyses ($p < 0.05$). Calibration graphs were drawn to show the differences between nomogram-predicted and actual survival rates of the colon cancer patients. Online version of the nomogram was established using “DynNom” and “shiny” R packages and deployed using shinyapps online website.

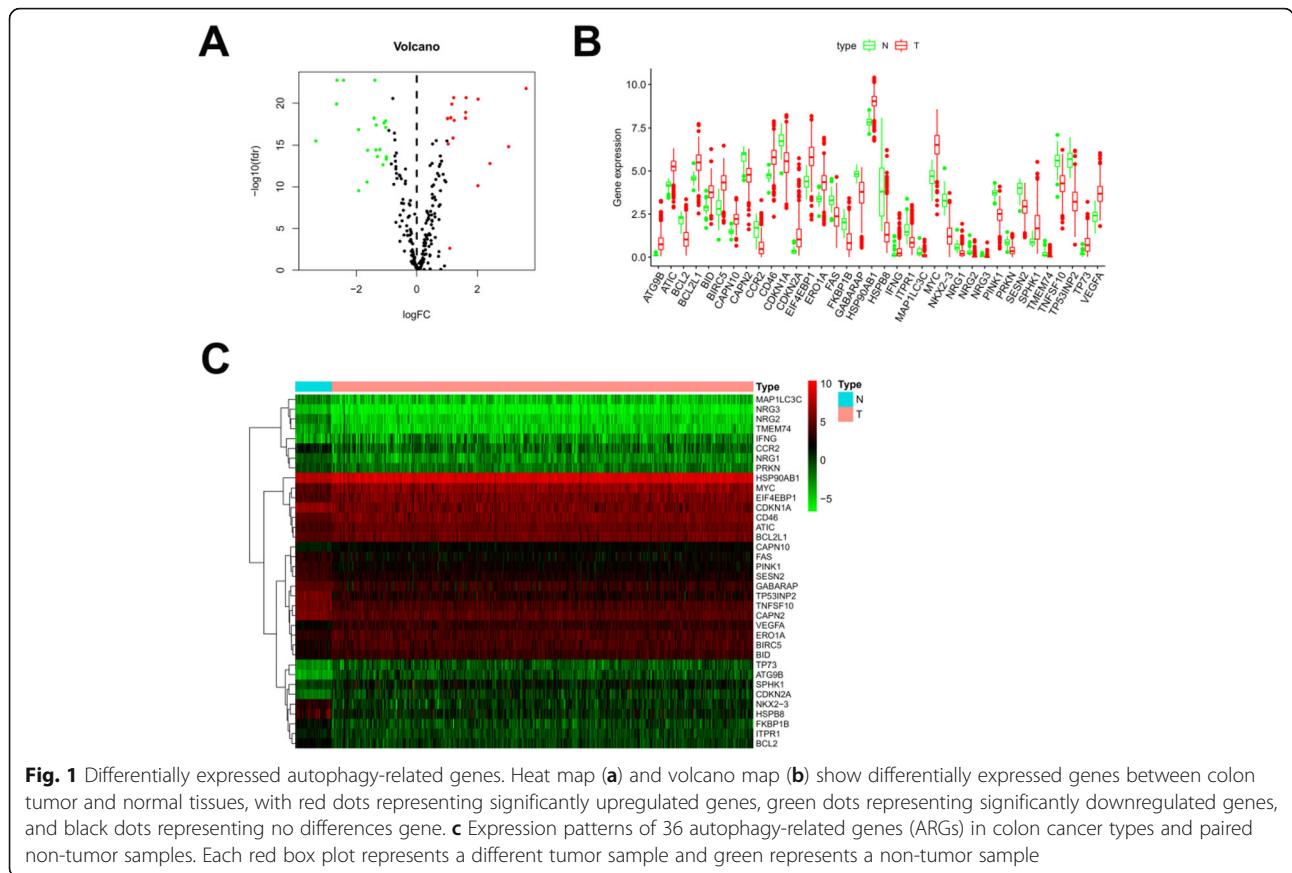
Statistical analysis

All statistical analyses were conducted using R programming language (v.4.0.2). Results with a p value < 0.05 were considered to be significant. Univariate Cox regression analysis was used to identify prognosis-related ARGs. Univariate and multivariate analyses were performed using the Cox proportional hazard model to identify factors that were independently related to prognosis of colon cancer patients. Survival curves were drawn using the Kaplan-Meier method and compared by a log-rank test.

Results

Differentially expressed ARGs

As shown in Fig. 1, 36 differentially expressed ARGs with a false discovery rate (FDR) < 0.05 and $|\log_{2}FC| > 1$ were identified from 232 ARGs. A volcano map (Fig. 1a), box-plots (Fig. 1b), and a heatmap (Fig. 1c) indicated that 20 ARGs (BCL2, CAPN2, CCR2, CDKN1A, FAS, FKBP1B, GABARAP, HSPB8, ITPR1, MAP1LC3C, NKX2-3, NRG1, NRG2, NRG3, PINK1, PRKN, SESN2, TMEM74, TNFSF10, and TP53INP2) were downregulated while 16 ARGs (ATG9B, ATIC, BCL2L1, BID, BIRC5, CAPN10, CD46, CDKN2A, EIF4EBP1, ERO1A, HSP90AB1, IFNG, MYC, SPHK1, TP73, and VEGFA) were overexpressed in colon tumor tissues. Functional enrichment analysis identified numerous GO and KEGG enrichment pathways (Fig. 2). The 36 genes were primarily related to the molecular functions of autophagy, process utilizing autophagic

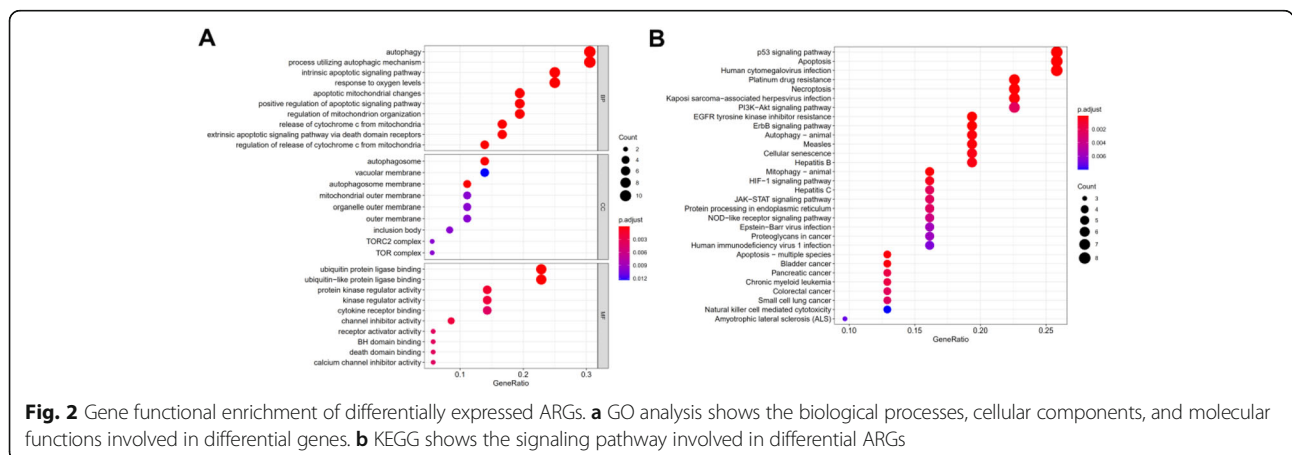


mechanism, and intrinsic apoptotic signaling pathways. As seen in Fig. 2a, these 36 genes are mainly related to molecular functions (MF) of autophagy, process utilizing autophagic mechanism, and intrinsic apoptotic signaling pathway, they are correlated with cellular components (CC) of autophagosome, vacuolar membrane, and autophagosome membrane, the genes are also involved in biological processes (BP) of ubiquitin-protein ligase binding, ubiquitin-like protein ligase binding, and protein kinase

regulator activity. These ARGs participate in the pathways of p53 signaling pathway, apoptosis, and human cytomegalovirus infection (Fig. 2b).

Prognosis-related ARGs

A forest map identified 20 ARGs that were associated with prognosis in colon cancer patients (Fig. 3a). Of these 20 prognosis-related genes, six genes were determined to be protective and 14 ARGs were associated



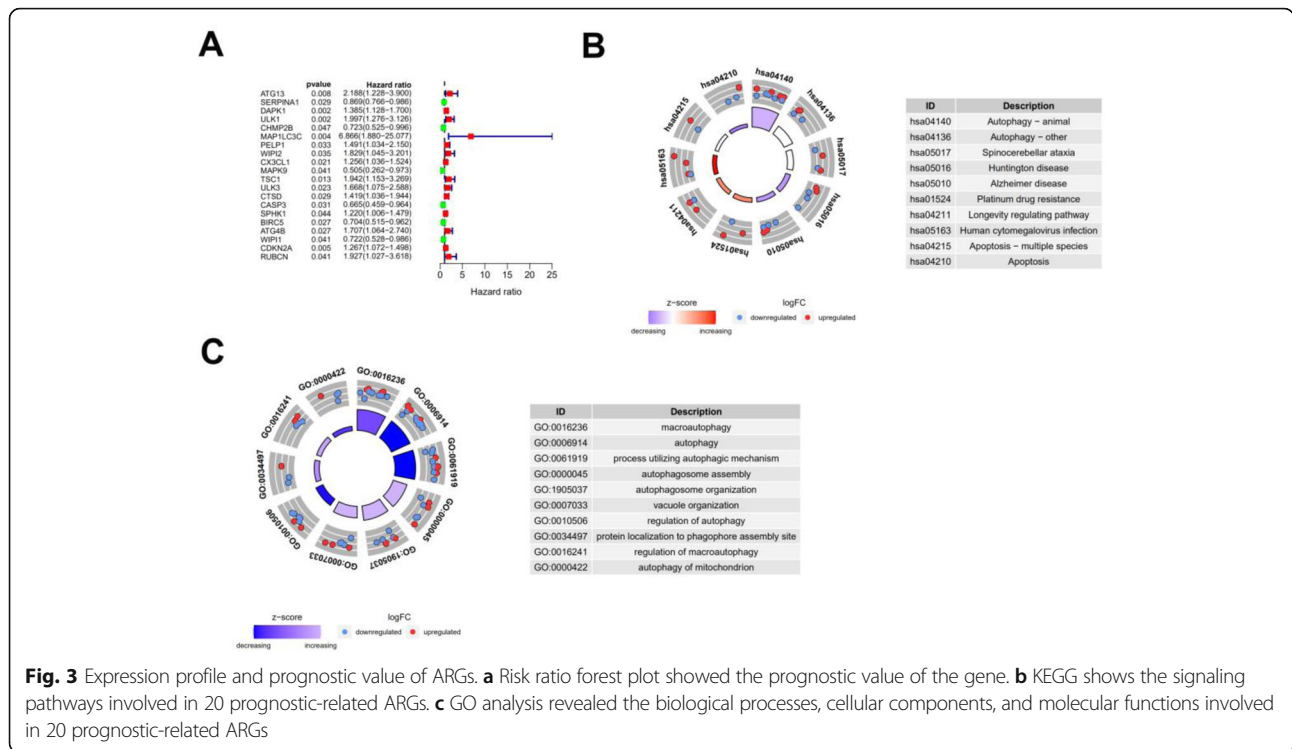


Fig. 3 Expression profile and prognostic value of ARGs. **a** Risk ratio forest plot showed the prognostic value of the gene. **b** KEGG shows the signaling pathways involved in 20 prognostic-related ARGs. **c** GO analysis revealed the biological processes, cellular components, and molecular functions involved in 20 prognostic-related ARGs

with increased risk. Results of the KEGG analysis indicated that prognostic ARGs were mainly involved in pathways of autophagy, spinocerebellar ataxia, and Huntington’s disease (Fig. 3b). Prognosis-related genes were correlated with macro-autophagy, autophagy, and process utilizing autophagic mechanism (Fig. 3c). Mutations in these 20 genes, examined using the cBioPortal website, showed that missense mutations, amplifications, and deep deletion were the most common mutation types (Fig. 4). Five ARGs (DAPK1, ULK1, PELP1, TSC1, and CASP3) had a mutation rate $\geq 3\%$, among which DAPK1 had the highest mutation rate.

Development of a prognosis model

Using multivariate analysis to develop a risk score for colon cancer patients, 8 ARGs were significantly related to prognosis. The risk score was defined as [Expression level of SERPINA1 $\times (-0.11979)$] + [Expression level of DAPK1 $\times (-0.29697)$] + [Expression level of MAP1LC3C $\times (1.50543)$] + [Expression level of MAPK9 $\times (-0.62080)$] + [Expression level of TSC1 $\times (-0.64199)$] + [Expression level of ULK3 $\times (-0.31259)$] + [Expression level of CASP3 $\times (-0.44136)$] + [Expression level of WIP1 $\times (-0.27200)$]. Based on the risk score, patients were divided into high-risk and low-risk groups using the median risk score as the

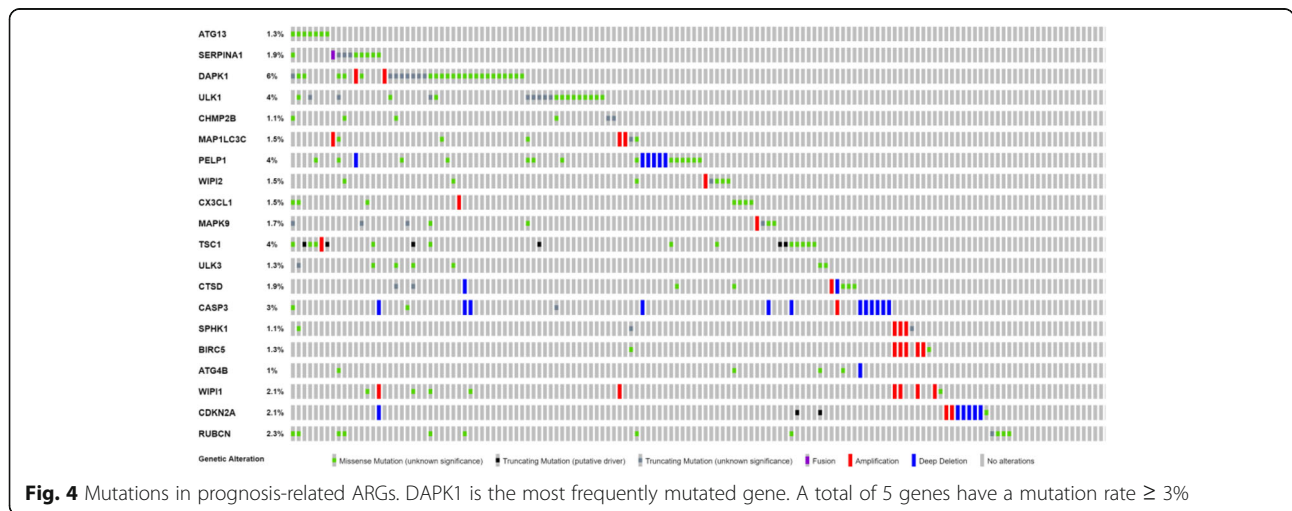


Fig. 4 Mutations in prognosis-related ARGs. DAPK1 is the most frequently mutated gene. A total of 5 genes have a mutation rate $\geq 3\%$

cut-off point between groups (Fig. 5a). Patients with higher risk scores were more likely to be deceased (Fig. 5b). A heatmap was used to show differences in expression for these 8 prognosis-related ARGs between groups (Fig. 5c).

Clinicopathologic characteristics of TCGA colon cancer patients were downloaded from TCGA database (Additional Table 1). Examination of the survival curves for the low-risk and high-risk patient groups, drawn using the Kaplan-Meier method (Fig. 5d), showed that high-risk patients had a significantly lower probability of survival ($p < 0.01$). Univariate and multivariate analyses were performed to identify prognosis-related factors in colon cancer patients (Fig. 6a and b). Factors with a p value < 0.05 in the univariate analysis were included in the multivariate analysis. Forest maps showed that age, pharmaceutical use, tumor invasion depth, lymph node metastasis, distant metastasis, and the risk score were still significant after multivariate analysis. Therefore, the risk score was independently associated with prognosis of patients [Hazard ratio (HR) = 1.537, 95% CI = 1.354-1.745, $p < 0.001$; Fig. 6b]. AUC of the ROC were used to test the prediction efficiency of the prognosis model (Fig. 6c). AUC of the risk score (0.701) was greater than that for any other clinicopathologic characteristics, including the American Joint Committee on Cancer stage, which showed that the risk score could be a reliable predictor of prognosis in colon cancer patients.

To better understand the influence of these factors on patient survival, a nomogram was drawn to predict 1-, 3-, and 5-year survival rates of colon cancer cases (Fig. 7a). The score obtained from the multivariate analysis was used to predict survival. Accordingly, if a 55-year old colon cancer patient with a tumor of T2N0M0 stage has a high calculated risk score, his or her estimated 5-year survival rate is 40 percent according to the predicted result of nomogram model. Moreover, calibration graphs depicting the differences between

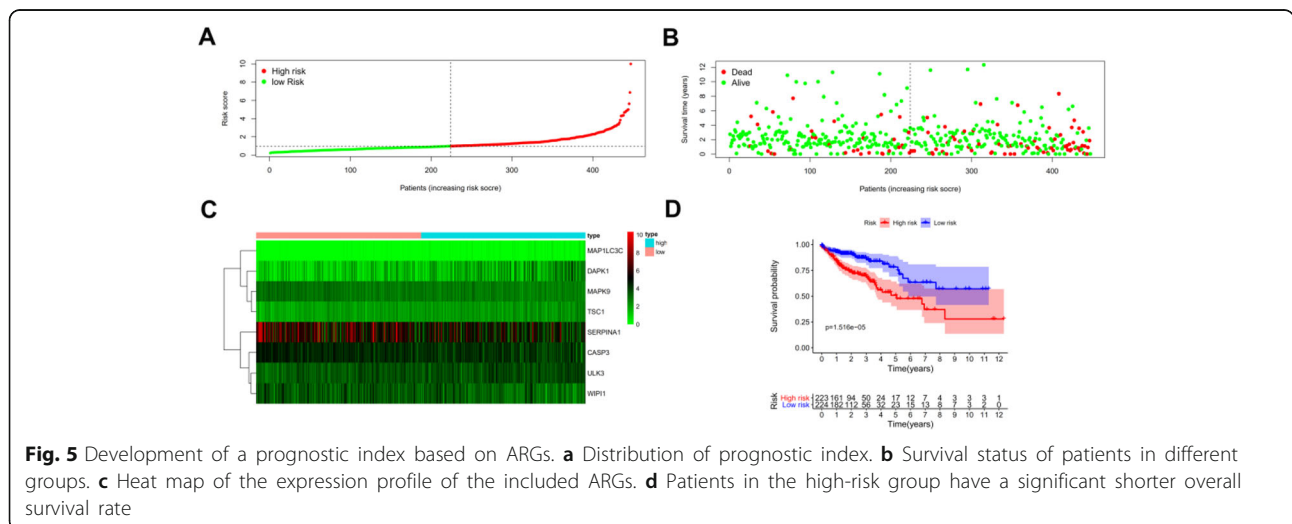
nomogram-predicted and actual survival rates of colon cancer patients showed that predicted 3- and 5-year survival rates were close to the actual survival rates (Fig. 7b and c), indicating that this nomogram model accurately predicted survival. Interestingly, the nomogram model was made into a web page at <https://doctorwang.shinyapps.io/DynNomapp>, which could be easily accessed using desktops, tablets, and smartphones (Additional Figure 2). The prognostic nomogram with an online webserver is more effective for providing accurate and individualized survival prediction in colon cancer patients.

The calculated risk score was associated with other clinicopathological characteristics, including tumor infiltration depth (Fig. 8a) and distant metastasis (Fig. 8b), suggesting that this model may also be predictive of tumor growth and metastasis.

Discussion

Autophagy has been proved to be associated with multiple types of cancer; however, the relationship between autophagy-related genes and prognosis in colon cancer patients remains largely unknown. To examine levels and patterns of expression of human ARGs in colon cancer, 36 differentially expressed ARGs were identified. GO and KEGG analyses were performed to explore relevant pathways and molecular biological functions. The most significant pathway identified in the KEGG analysis was the p53 signaling pathway. Mutations in the p53 gene occur in most types of malignancies [12], and the p53 signaling pathway plays an important role in cell cycle regulation, metabolism, development and aging, reproduction, and inhibition of tumor formation [13].

To explore the impact of ARGs on prognosis in colon cancer patients, 20 prognosis-related ARGs were identified using univariate Cox regression analysis. Eight of the 20 (40%) ARGs that remained significant after



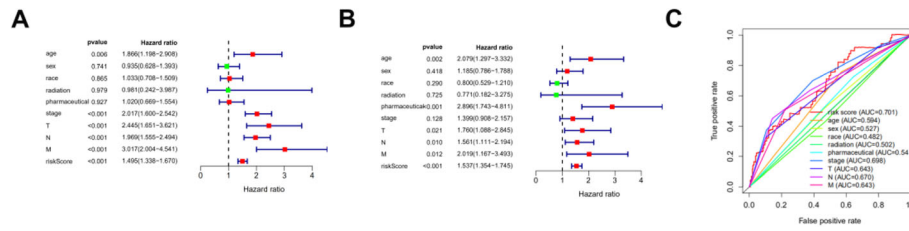


Fig. 6 Prognostic model based on ARGs shows good predictive performance. **(a)** Forest plot of univariate and **(b)** multivariate Cox regression analysis in colon cancer. **(c)** Survival-dependent receiver operating characteristic (ROC) curves validate the prognostic significance of ARGs-based prognostic indicators

multivariate analysis have been associated with prognosis in colon cancer or other malignant carcinomas. Gil et al. [14] observed that the expression of MAP1LC3C was downregulated in colorectal cancer tissues and was negatively associated with TNM stage. Yuan et al. [15] reported that downregulation of DAPK1 promotes chemoresistance and metastasis of colorectal cancer, while inhibition of DAPK1 promotes the epithelial-to-mesenchymal transition (EMT) of tumor stem cells. Hypermethylation of the MAPK9 promoter region affected the MAPK signaling pathway, focal adhesion, and Wnt signaling pathway in

colorectal cancer (CRC) [16]. Soo Jung Lee et al. [17] reported that genetic variation in the TSC1 gene may be useful as a biomarker for predicting patient outcomes after CRC resection surgery. High expression of SERP INA1 has been associated with advanced stage, lymph node metastasis, and poor prognosis of CRC patients, and may be useful as a prognostic marker and candidate therapeutic target for CRC [18]. Salemi et al. [19] observed overexpression of CASP3 in LNCaP and PC-3 prostate cancer cell lines. Upregulation of the ULK3 gene is known to occur in several tumor types, and ULK3 silencing

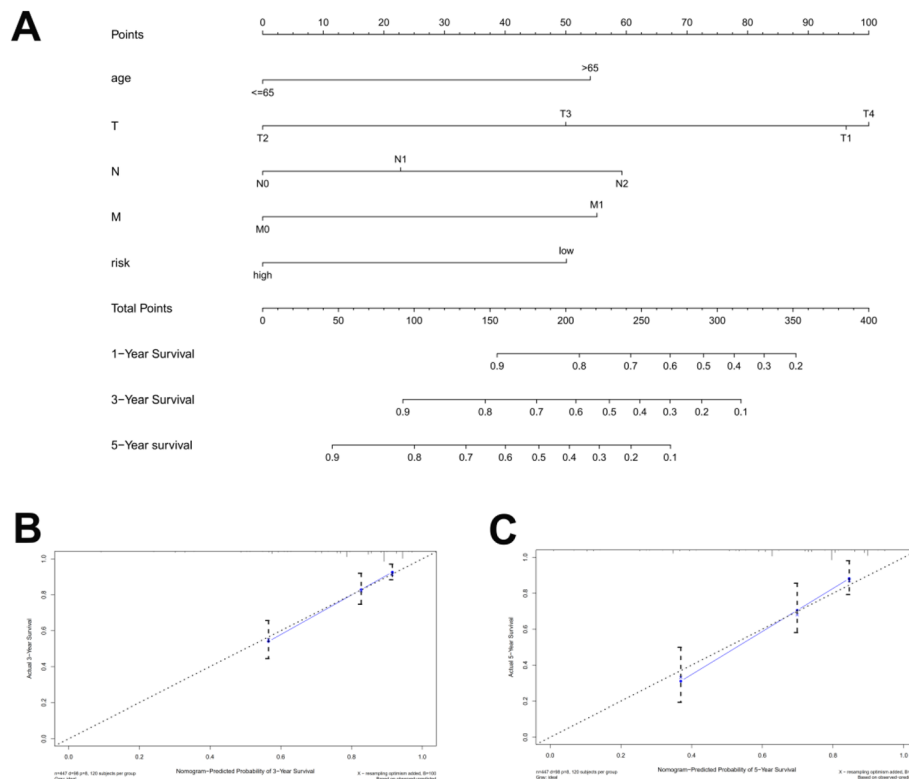
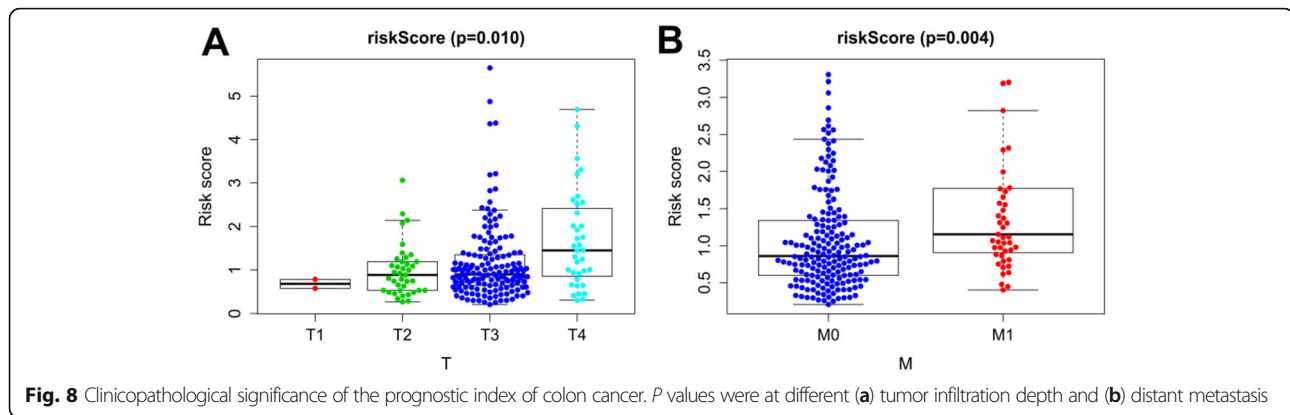


Fig. 7 Nomogram model **(a)** to predict 1-, 3-, and 5-year survival rates of colon cancer cases. Calibration graphs indicated that predicted 3- **(b)** and 5- **(c)** year survival rates were close to the actual survival rates



suppresses tumor progression. ULK3 connects two key signaling pathways in the transformation of normal fibroblasts into cancer-associated fibroblasts and, therefore, may represent a potential target for cancer therapy [20]. WIP1 has been proposed to be a new biomarker related to melanoma at both the gene and protein levels [21]. A previous research has documented that missense mutations constitutively activate oncoproteins [22] and that missense mutations are the most common type of mutation in ARGs associated with colon cancer.

Eight ARGs were used to predict patient prognosis and provide information for individualized treatment. To better understand the utility of clinical features and risk score on predicting outcomes, univariate and multivariate Cox analyses were performed. These analyses revealed that the calculated risk score is an independent predictor of patient prognosis. Because nomograms are widely used prediction tools in oncology, especially for cancer prognosis [23, 24], we developed a nomogram model to visualize the effects of clinical features and risk score on patients' 3- and 5-year survival probabilities. Calibration graphs verified that the nomogram had high prediction efficiency. Our nomogram of the online version provided more convenient and accurate prediction for colon cancer patients, it could be easily accessed by researchers and clinicians. As mentioned above, our risk score was associated with tumor infiltration depth and distant metastasis, suggesting that it may be related to the development and migration of colon cancer.

This study had several limitations. Subsequent molecular biological experiments are needed to further examine the function of ARGs in colon cancer development and to better understand carcinogenic mechanisms. Additional clinical cases will be required to maximize stability and the predictive ability of our established model.

In conclusion, our analysis of gene expression profiles and corresponding clinical characteristics identified prognosis-related ARGs in colon cancer. Genes associated with autophagy may represent new targets for developing improved treatment options and interventions

for patients with colon cancer. Based on the molecular characteristics of autophagy, we constructed a new risk scoring model, which can effectively evaluate the prognosis of colon cancer patients. Additional studies are needed to validate the findings of this study and provide a basis for individualized treatment.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12957-020-02061-w>.

Additional file 1: Table 1. Clinicopathologic characteristics of TCGA colon cancer patients.

Additional file 2: Figure 1. Flow chart of this study.

Additional file 3: Figure 2. Online version of the nomogram model.

Abbreviations

ARG: Autophagy-related gene; CEA: Carcinoembryonic antigen; GO: Gene ontology; KEGG: Kyoto encyclopedia of genes and genomes; ROC: Receiver operator characteristic curve; AUC: Area under curve; LogFC: Log fold change; FDR: False discover rate; CC: Cellular components; BP: Biological processes; MF: Molecular functions; CRC: Colorectal cancer

Acknowledgements

We would like to thank TopEdit (www.topedit.com) for English language editing of this manuscript.

Authors' contributions

Xu Wang is responsible for writing and submitting the papers; Yuanmin Xu is responsible for data collection and analysis; Ting Li and Bo Chen are responsible for the production of pictures; Wenqi Yang is responsible for the ideas and guidance. The authors have read and approved the final manuscript.

Funding

The Key Research and Development Plan Projects of Anhui Province (No. 201904a07020055).

Availability of data and materials

All the data and materials are available.

Ethics approval and consent to participate

There were no cell, tissue, or animal studies. No ethical requirements are involved.

Consent for publication

All authors agree to publish the paper.

Competing interests

The authors declare that they have no competing interests.

Received: 5 July 2020 Accepted: 21 October 2020

Published online: 30 October 2020

References

- 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.
- Doonan BB, Schaafsma E, Pinto JT, Wu JM, Hsieh TC. Application of open-access databases to determine functional connectivity between resveratrol-binding protein QR2 and colorectal carcinoma. *In Vitro Cell Dev Biol Anim*. 2017;53(7):575–8.
- Bhandari A, Woodhouse M, Gupta S. Colorectal cancer is a leading cause of cancer incidence and mortality among adults younger than 50 years in the USA: a SEER-based analysis with comparison to other young-onset cancers. *J Investig Med*. 2017;65(2):311–5.
- Russo AG, Andreano A, Sartore-Bianchi A, Mauri G, Decarli A, Siena S. Increased incidence of colon cancer among individuals younger than 50 years: a 17 years analysis from the cancer registry of the municipality of Milan, Italy. *Cancer Epidemiol*. 2019;60:134–40.
- Lenz HJ. Colon cancer stem cells: a new target in the war against cancer. *Gastrointest Cancer Res*. 2008;2(4):203–4.
- Yun CW, Lee SH. The roles of autophagy in cancer. *Int J Mol Sci*. 2018;19(11):3466.
- Levine B. Cell biology: autophagy and cancer. *Nature*. 2007;446(7137):745–7.
- Burada F, Nicoli ER, Ciurea ME, Uscatu DC, Ioana M, Gheonea DI. Autophagy in colorectal cancer: an important switch from physiology to pathology. *World J Gastrointest Oncol*. 2015;7(11):271–84.
- Schroll MM, Liu X, Herzog SK, Skube SB, Hummon AB. Nutrient restriction of glucose or serum results in similar proteomic expression changes in 3D colon cancer cell cultures. *Nutr Res*. 2016;36(10):1068–80.
- Mathew R, Karantza-Wadsworth V, White E. Role of autophagy in cancer. *Nat Rev Cancer*. 2007;7(12):961–7.
- Devenport SN, Shah YM. Functions and implications of autophagy in colon cancer. *Cells*. 2019;8(11):1349.
- Dulak AM, Stojanov P, Peng S, et al. Exome and whole-genome sequencing of esophageal adenocarcinoma identifies recurrent driver events and mutational complexity. *Nat Genet*. 2013;45(5):478–86.
- Hu W, Feng Z, Teresky AK, Levine AJ. p53 regulates maternal reproduction through LIF. *Nature*. 2007;450(7170):721–4.
- Gil J, Ramsey D, Pawlowski P, et al. The influence of tumor microenvironment on ATG4D gene expression in colorectal cancer patients. *Med Oncol*. 2018;35(12):159.
- Yuan W, Ji J, Shu Y, et al. Downregulation of DAPK1 promotes the stemness of cancer stem cells and EMT process by activating ZEB1 in colorectal cancer. *J Mol Med (Berl)*. 2019;97(1):89–102.
- Tian X, Sun D, Zhao S, Xiong H, Fang J. Screening of potential diagnostic markers and therapeutic targets against colorectal cancer. *Onco Targets Ther*. 2015;8:1691–9.
- Lee SJ, Kang BW, Chae YS, et al. Genetic variations in STK11, PRKAA1, and TSC1 associated with prognosis for patients with colorectal cancer. *Ann Surg Oncol*. 2014;21(Suppl 4):S634–9.
- Kwon CH, Park HJ, Choi JH, et al. Snail and serpinA1 promote tumor progression and predict prognosis in colorectal cancer. *Oncotarget*. 2015;6(24):20312–26.
- Salemi M, Condorelli RA, La Vignera S, et al. PARP-1 and CASP3 genes are up-regulated in LNCaP and PC-3 prostate cancer cell lines. *Hum Cell*. 2014;27(4):172–5.
- Goruppi S, Procopio MG, Jo S, Clocchiatti A, Neel V, Dotto GP. The ULK3 kinase is critical for convergent control of cancer-associated fibroblast activation by CSL and GLI. *Cell Rep*. 2017;20(10):2468–79.
- D'Arcangelo D, Giampietri C, Muscio M, Scatozza F, Facchiano F, Facchiano A. WIP1, BAG1, and PEX3 autophagy-related genes are relevant melanoma markers. *Oxid Med Cell Longev*. 2018;2018:1471682.
- Kamburov A, Lawrence MS, Polak P, et al. Comprehensive assessment of cancer missense mutation clustering in protein structures. *Proc Natl Acad Sci U S A*. 2015;112(40):E5486–95.
- Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol*. 2008;26(8):1364–70.
- Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol*. 2015;16(4):e173–80.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

