Prognostic model and immune-infiltrating cell landscape based on differentially expressed autophagy-related genes in TP53-mutated multiple myeloma

**Type**
Research paper

**Keywords**
prognosis, multiple myeloma, autophagy, risk signature, TP53 mutation, immune-infiltrating cell

**Abstract**

Introduction
Autophagy functions as a prosurvival mechanism in multiple myeloma (MM). The objective of this research was to establish an autophagy-related gene (ARG) signature for predicting the survival outcomes of MM patients with TP53 mutations.

Material and methods
Information about MM patients with TP53 mutations was downloaded from Gene Expression Omnibus (GEO) database. Cox proportional hazard regression was employed to determine the independent prognostic ARG and construct a risk signature. Time-dependent receiver-operating characteristic (t ROC) curve was used to explore the predictive accuracy of the prognostic model. A nomogram was constructed to give a more precise prediction of the probability of 5-year, 8-year and 10-year overall survival (OS). In addition, we utilized the CIBERSORT algorithm to explore the distribution difference of 22 immune-infiltrating cells.

Results
Three differentially expressed ARGs (CASP8, MAPK8, RB1CC1) were finally incorporated to construct the risk model. Area under the curve (AUC) of corresponding tROC curve for 5-year, 8-year and 10-year OS were 0.735, 0.686 and 0.662, respectively. MM patients were categorized into high and low-risk group in accordance with the median threshold value (-1.724549). ARG-based risk score model was an independent prognostic element correlated with OS, giving an hazard ratio (HR) of 3.29 (95%CI 2.35-4.60, P<0.001). 13 immune infiltrating cells were found to have distribution differences between the two groups.

Conclusions
We established a three-ARGs risk signature which manifested an independent prognostic factor. The nomogram was testified to perform well in forecasting the long-term survival of TP53-mutated MM patients.
Dear distinguished editors,
This is about our revised manuscript (AMS-13124-2021-01) entitled "Prognostic model and immune-infiltrating cell landscape based on differentially expressed autophagy-related genes in TP53-mutated multiple myeloma" to Archives of Medical Science.

First, I would like to extend my sincere gratitude for the instructive comments from the conscientious editors and the reviewers.

A revised version of the manuscript (with track changes marked in red) has also been reuploaded in the submission system.

In accordance with your requirements and comments, we have made a point-by-point response to all the reviewers and editors, and the modifications are summarized as follows.

Review 1:
In the current study, Zheng et al established a prognostic model and immune-infiltrating cell landscape based on differently expressed autophagy-related genes in TP53-mutated multiple myeloma, and proposed a three autophagy-related gene risk signature which was testified to perform well in forecasting the long-term survival of multiple myeloma patients with TP53-mutations. In addition, they found thirteen kinds of immune infiltrating cells were distributed differently between the high and low-risk group according to the prognostic model. Their findings are interesting and have clinical merit to MM management.

There were still two minor points remained to be addressed or explained:

1. Table 3 listed the distribution differences of 22 immune-infiltrating cells. The increase or decrease of immune cells need to be marked in the manuscript.

Author response to reviewer 1:
Thanks for your kind suggestions. We have specified and added the increase or decrease of immune cells in our manuscript as required.
The newly-added contents are:
“The abundance of CD8+T cells, gamma delta T cells, activated NK cells, 
activated dendritic cells, monocytes, resting mast cells, and macrophages-
M0,M1 were observed lower in the high-risk group than the low-risk 
group(all p-value<0.05), whereas the fraction of memory B cells, plasma 
cells, naïve CD4+T cells, resting NK cells, and activated mast cells were 
higher in the high-risk group(all p-value<0.05). The results indicated that the 
proportion of different immune-infiltrating cells was closely associated with 
the aggressiveness and risk stratification and DEARG-based risk signature 
might be correlated with the immune microenvironments of TP53-mutated MM.”

2.The ROC value of risk model is not so prominent. The authors are thus 
recommended to further compare or explain the difference of their model with 
traditional ISS staging system in multiple myeloma.

Author response to reviewer1: 
Thanks for your comments. Just as you have mentioned, we also noticed that the 
ROC value of our established model is not so prominent as expected. But our 
study was only focused on TP53-mutated MM patients. Therefore, we consider it 
unsuitable to compare our established model with traditional ISS staging system 
or Durie-Salmon staging system in MM.

Review 2: 
1.Results, section 3.5: differences of cell populations should be explained clearly 
in the text, for example ..... monocytes infiltration is higher in low risk cases...

Author response to reviewer2: 
Thanks for your instructions. Just as reviewer 1 have pointed out the same 
question, we have specified and added the increase or decrease of immune cells 
in our manuscript.
The newly-added contents are:
“The abundance of CD8+T cells, gamma delta T cells, activated NK cells, activated dendritic cells, monocytes, resting mast cells, and macrophages-M0,M1 were observed lower in the high-risk group than the low-risk group (all p-value<0.05), whereas the fraction of memory B cells, plasma cells, naïve CD4+T cells, resting NK cells, and activated mast cells were higher in the high-risk group (all p-value<0.05). The results indicated that the proportion of different immune-infiltrating cells was closely associated with the aggressiveness and risk stratification and DEARG-based risk signature might be correlated with the immune microenvironments of TP53-mutated MM.”

2. Actually, I must say and share with you that, a constant finding (as my observation) in bone marrow microscopy of myeloma patients is this: when monocytes are present, it is good, patients do well, progression is slow. But when monocytes are dissappeared, things are not good, disease is (or will be) highly aggressive. So, I am highly impressed by your results.

**Author response to reviewer2:**
Thanks very much for your comments. We are also happy to hear that our study results are consistent with your observations. And to some extent, your observation corroborated our results.

**Review 3:**
I think that the Survival analysis and establishment of prognostic model should be expanded better and the CIBERSORT estimation of immune infiltration cells should be validated in vitro, for exemple. It is difficult understand, biologically, how with the 3 genes (DEARGs) you arrive in the risk score and in the prognostic model.

**Author response to reviewer3:**
Thank you for your comments. As I have mentioned in the Materials and Methods Section, the process of identifying and incorporating the DEARG in our prognostic
model has already been elucidated in the section “3. Survival analysis and establishment of prognostic model”.

Here we will give a detailed description of the process: “Log-rank test was used to evaluate DEARG which might be associated with the prognosis, with p-value<0.05 deemed statistically significant. Then DEARGs with potential prognostic value were initially selected by means of univariate Cox regression model. Least absolute shrinkage and selection operator (LASSO) regression was then utilized to eliminate false positive DEARGs because of over-fitting. DEARGs with p-value<0.05 in the univariate results were integrated into the multivariate analysis to determine the independent prognostic factors associated with OS and then construct a risk signature.” “The risk score for each individual patient was quantified by the following formula: risk scores=$\sum_{i=1,2,...,n}\beta(DEARGi) \times \text{Exp}(i)$, where $\beta$ represents the regression coefficient for each DEARG derived from the multivariate Cox regression and Exp indicates the relative expression levels of each DEARG standardized by Z-score. Survival difference between the above two groups were also evaluated by Kaplan-Meier curve and then the log-rank test. Univariate and multivariate Cox regression were further carried out to determine whether the DEARG-based risk score could be an independent prognostic factor in TP53-mutated MM patients.”

Section Editor recommendation
Accept without changes

Author response to Section editor:
Thanks very much for your suggestions.

Furthermore, I have cited the following relevant articles published in last 2-3 years in order to build the impact factor.

The newly-cited references in our manuscript which were published in *Archives of Medical Science* are as follows.


The corresponding newly-added contents are as follows. Also, induction of autophagy forcefully strengthened chemoresistance to gemcitabine in bladder carcinoma (Reference 29).

Our study revealed that immune cells with tumor-killing effect in high-risk group constituted smaller proportion than their counterparts in low-risk group, which is consistent with a previous research on gastric cancer (Reference 52).

Thanks again for your generous assistance. We hope that the revised version can meet the requirements. I am looking forward to hearing from you.

Cordially Yours
Yan-Hua Zheng
Prognostic model and immune-infiltrating cell landscape based on differentially expressed autophagy-related genes in TP53-mutated multiple myeloma

Running head: ARGs risk signature in TP53-mutated MM

Yan-Hua Zheng¹, Hong-Yuan Shen¹, Xiang Chen², Juan Feng¹, Guang-Xun Gao¹*

Yan-Hua Zheng and Hong-Yuan Shen contributed equally to this article as co-first authors

Affiliations:
1 Department of Hematology, Xijing Hospital, Fourth Military Medical University, Xi’an, Shaanxi, 710032, China
2 Department of Pharmacy, Daping Hospital, Third Military Medical University, Chongqing, 400042, China

*Corresponding author: Professor Guang-Xun Gao, Department of Hematology, Xijing Hospital, Fourth Military Medical University, 127 Chang’le West Road, Xi’an, Shaanxi, 710032, China.
E-mail: gaoguangxun@fmmu.edu.cn Tel: 86-29-84775199
Abstract

Introduction: Autophagy functions as a prosurvival mechanism in multiple myeloma (MM). The objective of this research was to establish an autophagy-related gene (ARG) signature for predicting the survival outcomes of MM patients with TP53 mutations.

Material and Methods:

Information about MM patients with TP53 mutations was downloaded from Gene Expression Omnibus (GEO) database. Cox proportional hazard regression was employed to determine the independent prognostic ARG and construct a risk signature. Time-dependent receiver-operating characteristic (t ROC) curve was used to explore the predictive accuracy of the prognostic model. A nomogram was constructed to give a more precise prediction of the probability of 5-year, 8-year and 10-year overall survival (OS). In addition, we utilized the CIBERSORT algorithm to explore the distribution difference of 22 immune-infiltrating cells.

Results:

Three differentially expressed ARGs (CASP8, MAPK8, RB1CC1) were finally incorporated to construct the risk model. Area under the curve (AUC) of corresponding tROC curve for 5-year, 8-year and 10-year OS were 0.735, 0.686 and 0.662, respectively. MM patients were categorized into high and low-risk group in accordance with the median threshold value (-1.724549). ARG-based risk score model was an independent prognostic element correlated with OS, giving an hazard ratio (HR) of 3.29 (95%CI 2.35-4.60, P<0.001). 13 immune infiltrating cells were found to have distribution differences between the two groups.

Conclusions:
We established a three-ARGs risk signature which manifested an independent prognostic factor. The nomogram was testified to perform well in forecasting the long-term survival of TP53-mutated MM patients.

**Keywords:** multiple myeloma, autophagy, prognosis, risk signature, TP53 mutation, immune-infiltrating cell
1. Introduction

Multiple myeloma (MM) represents the second most common hematological malignancy. It is characterized by the proliferative disturbance of plasma cells within the bone marrow, resulting in excessive accumulation of monoclonal immunoglobulins in the blood or urine[1]. MM is often clinically manifested by “CRAB” symptoms (hypercalcemia, renal dysfunction, anemia and bone lesions)[2]. Despite therapeutic improvements with wide combinative application of proteasome inhibitors, immunomodulatory agents, and monoclonal antibodies[3], MM remains an incurable disease with high relapse rate and relatively poor prognosis[4].

TP53 gene, a well-known tumor suppressive gene, is situated at the chromosome 17p13.1 and codes for p53 protein. P53 protein is described as "the guardian of genome" for its pivotal role in maintaining genomic integrity and cellular homeostasis[5, 6]. TP53 mutation is an adverse prognostic factor in various cancers, including solid tumors and hematological malignancies such as acute myelogenous leukemia (AML), acute lymphocytic leukemia (ALL), chronic lymphocytic leukemia (CLL), myelodysplastic syndrome (MDS) and MM[7, 8]. Compared with MM patients without TP53 mutation, TP53-mutated patients have shorter overall survival (OS) and bleaker prognosis. TP53 mutation is exclusively correlated with del(17p) in MM[9-13].

Autophagy is a highly-conserved multi-step metabolic process in which cellular proteins and organelles are engulfed by autophagosomes and then transported to lysosomes for degradation[14]. Tightly controlled and modulated by a cluster of autophagy-related genes (ARGs), it can be stimulated in adverse circumstances including nutrients deficiency, hypoxia and DNA damage. Autophagy plays a critical role in -modulating cellular self-clearance, providing energy and maintaining the homeostasis and survival by re-utilizing the components...
such as amino acids, fatty acids, and nucleotides[15]. Autophagy constitutes a double-edged sword in tumorigenesis and progression. Whether autophagy promotes or represses cancer depends upon the type and stage of specific cancer, which renders targeting autophagy in cancer treatment controversial[16, 17]. In MM cells, proteasome inhibition brings about the accumulation of misfolded proteins, thus instigating endoplasmic reticulum(ER) overload and stress through unfolded protein response. Autophagy functions as a prosurvival mechanism through which MM cells develop resistance to proteasome inhibitors and evade from excessive accumulation of toxic proteins[18]. So, targeting autophagy might be a promising therapeutic strategy to prompt cell apoptosis and restore drug sensitivity, thus augmenting the efficacy of conventional chemotherapy[19, 20].

The above results corroborated the vital role of autophagy in MM and ARGs might carry clinical implications as potential prognostic biomarkers. In contrast with a single gene, prognostic model incorporating multiple ARGs may greatly enhance the predictive performance. Up till now, there existed few studies integrating a set of ARG expression signature for predicting the survival outcomes of MM patients with TP53 mutations. The objective of this study was to establish a more accurate predictive model with ARGs signature. TP53-mutated MM patients were identified from the Gene Expression Omnibus (GEO) database and were subdivided into high and low risk groups in accordance with the median predictive value. By applying the CIBERSORT method, we also explored the distribution difference of 22 immune-infiltrating cell subsets within the bone marrow microenvironment between the two groups.

2. Materials and Methods
2.1. Patient information and datasets processing

GSE136400 datasets were acquired from the GEO database (GEO, https://www.ncbi.nlm.nih.gov/geo/) for the clinical characteristics, gene expression profile and OS information of MM patients.

ARGs were obtained by retrieving the Human Autophagy Database (HADb, http://autophagy.lu/clustering/index.html).

Differentially expressed genes (DEGs) in MM patients with or without TP53 mutations were identified and analyzed by means of linear models for microarray data (LIMMA, the “limma” package of R software)[21]. DEGs, including both significantly up-regulated and down-regulated genes, were determined by utilizing the Wilcoxon signed-rank test. The cut-off value was defined as the false discovery rate (FDR) < 0.05. And the status of DEGs was demonstrated in volcano plot and heatmap.

2.2. Enrichment analysis of differentially expressed autophagy-related genes (DEARGs)

We speculated that DEGs might be intersected with ARGs. We defined genes which overlapped on both databases as overlapping candidate genes (OCGs) or differentially expressed autophagy-related genes (DEARGs). Subsequently, both gene ontology (GO) functional enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways analysis were conducted on DEARGs via utilizing the “clusterProfiler” package in R software, with an adjusted p-value<0.05 regarded statistically significant[22]. Also, gene sets with an FDR score<0.05 were regarded noticeably enriched[23].

2.3. Survival analysis and establishment of prognostic model
OS was calculated from the date of initial diagnosis until death from all cause or the last follow-up, whichever came first. Kaplan-Meier survival curves for OS were plotted with the purpose of comparing each potential highly and lowly-expressed DEARGs. Log-rank test was used to evaluate DEARG which might be associated with the prognosis, with p-value<0.05 deemed statistically significant. Then DEARGs with potential prognostic value were initially selected by means of univariate Cox regression model. Least absolute shrinkage and selection operator (LASSO) regression was then utilized to eliminate false positive DEARGs because of over-fitting. DEARGs with p-value<0.05 in the univariate results were integrated into the multivariate analysis to determine the independent prognostic factors associated with OS and then construct a risk signature.

The risk score for each individual patient was quantified by the following formula: risk scores=$\sum_{i=1,2...n} \beta(DEARGi) \times Exp(i)$, where $\beta$ represents the regression coefficient for each DEARG derived from the multivariate Cox regression and Exp indicates the relative expression levels of each DEARG standardized by Z-score. Patients (in the training set) were categorized into high and low-risk group using the median risk score as the threshold. A high-risk score represented bleaker prognosis than low-risk score. Survival difference between the above two groups were also evaluated by Kaplan-Meier curve and then the log-rank test. Univariate and multivariate Cox regression were further carried out to determine whether the DEARG-based risk score could be an independent prognostic factor in TP53-mutated MM patients. We adopted time-dependent receiver-operating characteristic (tROC) analysis to explore the predictive accuracy of the prognostic model, which could be quantified by the area under the ROC curve (AUC).
Eventually, the “rms” package of R software was employed to construct the nomogram, which incorporated all independent prognostic parameters (perhaps including DEARG-based risk model and other clinical factors), to give a more precise prediction of 5-year, 8-year and 10-year OS probability. Then, the concordance between actually-observed and predicted survival was assessed through calibration curves, in which the 45° line denoted the best predictive performance.

2.4. CIBERSORT estimation of immune infiltration cells

CIBERSORT (Cell type Identification By Estimating Relative Subsets Of RNA Transcripts) algorithm, which can be available through web portal (http://cibersort.stanford.edu/), is based on a machine-learning approach named support vector regression. The CIBERSORT algorithm was employed to calculate the proportions of 22 kinds of immune infiltrating cells including B cells, T cells, natural killer cells, macrophages, dendritic cells and so on[24]. For each TP53-mutated patient, the final CIBERSORT estimated-results were standardized and the proportions of 22 kinds of immune-infiltrating cells summed up to 1. We then evaluated the proportion differences between high and low-risk group on the basis of DEARG-based score[25].

2.5. Statistical analysis

R software (version 3.5.1) was utilized to conduct all statistical analyses. All tests were two-sided with p-value < 0.05 deemed significant difference.
3. Results

3.1. Identification of DEARGs and functional enrichment analysis.

Information about 557 multiple myeloma patients with TP53 mutations and 400 cases without TP53 mutations was downloaded from GSE136400 datasets of GEO. Our study was mainly focused on TP53-mutated patients. The patient demographic and clinical characteristics were revealed in Table 1.

Based on the criteria for FDR<0.05 by using “limma” package, we identified a set of 3329 DEGs between the TP53-mutated patients and TP53-unmutated patients, containing 2745 up-regulated and 584 down-regulated DEGs. The volcano plot and the heatmap of DEGs were demonstrated in Figure 1A and Figure 1B respectively.

We extracted altogether 222 ARGs from HADb database. Then 3329 DEGs were intersected with 222 ARGs and thus we obtained 51 OCGs called DEARGs. The Venn diagram of DEGs and ARGs was shown in Figure 1C.

KEGG and GO enrichment analysis were conducted with the aim of providing a panoramic view of biological function of 51 DEARGs and then were demonstrated in Figure 2A and Figure 2B. The more genes were enriched in the corresponding terms, the darker the color was. KEGG analysis indicated the signaling pathways which were implicated in autophagy, mitophagy and so on. GO analysis showed that regulation of autophagy was the main biological processes and molecular functions of DEARGs.

3.2. Screening and verification of prognosis-associated DEARGs by survival analysis.

In altogether 557 TP53-mutated MM patients, 90 cases lacked survival information. Finally, 467 cases were included for further analysis. In order to identify the DEARGs
associated with the prognosis, Kaplan-Meier curves for OS, log-rank test and univariate Cox regression were performed by comparing the highly and lowly-expressed DEARGs. A total of 9 prognosis-associated DEARGs was selected, including \textit{ATG2A, ATG2B, BIRC6, CASP8, CCL2, CFLAR, MAPK1, MAPK8, RB1CC1}. The above 9 DEARGs with p-value<0.05 in the univariate analysis were integrated into LASSO regression model and multivariate Cox regression model. Finally, we obtained 3 DEARGs (\textit{CASP8, MAPK8, RB1CC1}), which were independent risk factors. The survival curves for the above-mentioned 3 DEARGs were demonstrated in Figure 3.

3.3. Establishment and estimation of risk score model for predicting OS.

According to the multivariate Cox coefficients, the 3 DEARG-based risk score was constructed. The formula is: \text{risk score} = \text{Exp(CASP8)}*0.73254+\text{Exp(MAPK1)}*(-0.42603)+\text{Exp(RB1CC1)}*(-0.36719), in which Exp denoted the expression level of each DEARG.

We conducted the tROC analysis to determine the sensitivity and specificity of the predictive model. We achieved the results that AUC of corresponding ROC curve for 5-year, 8-year and 10-year OS were 0.735, 0.686 and 0.662, respectively (Figure 4A).

Then we calculated the risk score of individual patients, obtaining the median cut-off point of -1.724549. We categorized the patients into high-risk group(n=233) and low-risk group(n=234) according to the median threshold value. As demonstrated in Kaplan-Meier survival curve (Figure 4B), the high-risk group had a bleaker prognosis compared with the low-risk group(p<0.0001). Figure 4 indicated that DEARG-based risk score performed well in OS prediction of TP53-mutated MM patients.
3.4. Construction of the nomogram prognostic model

We further performed the univariate and multivariate Cox regression to analyze the correlation between the risk-score model and age, gender and International Staging System (ISS). As is revealed in Table 2, DEARG-based risk score model (HR 3.29, 95%CI 2.35-4.60) and ISS stage (HR=1.90, 95%CI 1.57-2.30) are independent prognostic factors correlated with OS (p<0.001). With the purpose of establishing a more accurate prognostic model, we constructed a nomogram which incorporated ISS stage and risk score to forecast the 5-year, 8-year and 10-year OS of TP53-mutated patients (Figure 5A). However, the predictive outcome of 5-year OS was not achieved. The calibration plot verified that the predictive performance of the nomogram for 8-year and 10-year OS agreed roughly with the actual outcome (Figure 5B).

3.5. Subpopulations of immune-infiltrating cells by CIBERSORT estimation

The proportions of immune-infiltrating cells were estimated by applying CIBERSORT algorithm and the LM22 gene signature with 1000 permutations, which could discriminate the phenotypes of 22 immune-infiltrating cells both sensitively and specifically[26]. Immune cell infiltration landscape was demonstrated in Figure 6. Then we analyzed the distribution differences between high-risk group and low-risk group among the TP53-mutated MM patients. In altogether 22 immune-infiltrating cell subtypes, statistical differences between the high and low risk group were found in 13 immune cells, including plasma cells, monocytes, resting mast cells, activated NK cells, activated dendritic cells, resting NK cells, memory B cells, activated mast cells, naïve CD4+T cells, follicular helper T cells, gamma delta T cells, macrophages-M0, CD8+T cells (Table 3). The abundance of CD8+T cells, gamma delta T cells, activated NK cells,
activated dendritic cells, monocytes, resting mast cells, and macrophages-M0,M1 were observed lower in the high-risk group than the low-risk group (all p-value < 0.05), whereas the fraction of memory B cells, plasma cells, naïve CD4+ T cells, resting NK cells, and activated mast cells were higher in the high-risk group (all p-value < 0.05). The results indicated that the proportion of different immune-infiltrating cells was closely associated with the aggressiveness and risk stratification and DEARG-based risk signature might be correlated with the immune microenvironments of TP53-mutated MM.

Discussion

Autophagy, a complicated multi-step self-digestion process, is regulated by a multiple of ARGs to guarantee homeostasis, energy supply and re-utilization[27]. Higher level of basal autophagy is often observed in MM cells and autophagy is vital for MM cell survival. MM cell apoptosis can be induced by autophagy disturbance via BECLIN-1 knockdown or pharmacologic repression with chloroquine or 3-methyladenine[28]. Our team revealed that autophagy inhibition by pharmacological methods augmented apoptosis in DNA-damaged MM cells and knockdown of Beclin 1 or ATG5 resensitizes MM cells to apoptosis induced by DNA-damaging agents[19]. Also, induction of autophagy forcefully strengthened chemoresistance to gemcitabine in bladder carcinoma[29].

Recently, many risk models based upon ARG signatures were established to forecast the survival outcomes of patients with non-small cell lung cancer[30], breast cancer[31], clear-cell renal cell carcinoma[32, 33], colorectal cancer[34], serous ovarian cancer[35], glioblastoma multiforme[36], prostate cancer[37] and bladder cancer[38]. Up till now, the prognostic effects of ARGs in TP53-mutated MM have not been comprehensively investigated. Therefore, our
study screened and finally identified three ARGs (CASP8, MAPK8, RB1CC1) to establish a risk model and predict the patient’s survival.

CASP8 gene is located on chromosome 2q33-34 and encodes caspase-8, which is a canonical cysteine protease for the initiation and the execution of cell apoptosis. Caspase-8 serves as a key component of death receptor-induced programmed cell death and is regarded a tumor suppressor[39]. Previous studies demonstrated activated caspases are also implicated in autophagy inhibition via cleaving autophagy-related proteins (Beclin-1, Atg5, and p62)[40, 41]. MAPK8(Mitogen-activated protein kinase 8) encodes Jun N-terminal kinase-1 (JNK1) and is the hallmark of the famous MAPK signaling pathway, which is involved in DNA damage response, autophagy, tumorigenesis and progression. RB1CC1 (Retinoblastoma coiled coil protein 1), also named FIP200 (FAK family-interacting protein of 200 kDa), serves as a constituent of the ULK1-ATG13-RB1CC1 or RB1CC1-ATG101 complex and plays an essential role in autophagosome formation. RB1CC1 is situated both in the nucleus and in the cytoplasm. RB1CC1 modulates intracellular signaling pathways through interacting with TSC1, p53, and PIASy, thus affecting cell cycle, cell proliferation and differentiation[42, 43].

Traditional ISS only incorporated laboratory parameters such as serum albumin, lactate dehydrogenase, beta-2-microglobulin. Cytogenetic abnormalities were introduced into Mayo clinic risk stratification for multiple myeloma (mSMART), which integrated del(17p),gain(1q),t(4;14),t(14;16),t(14;20)[2]. Both ISS and mSMART exhibited limited performance in MM risk stratification. We performed multivariate analysis to verify whether this three-DEARG-based risk model for survival prediction is independent of other prognostic covariates. Furthermore, we initiated a nomogram to forecast the 5-year, 8-year and 10-year OS
of individual patient. Nomogram has been widely used to quantitatively determine the clinical outcome on an individual level by combining each independent factor.

Immune-infiltrating cells constitute an integral part of the tumor microenvironment. By applying a machine-learning approach termed support vector regression, CIBERSORT is considered the most accurate algorithm which allows for highly sensitive and specific discrimination and quantification of the proportions of 22 human immune infiltrating cells. CIBERSORT has also been used in construction of immune-related risk model for several cancer types. Recently, accumulating researches have investigated the effects of the immune microenvironment on various cancers, including lung squamous cell carcinoma[44], hepatocellular carcinoma[45, 46], cholangiocarcinoma[47], colon cancer[48, 49], breast cancer[50] and cutaneous melanoma[51], which indicated that the infiltration of different types of immune cells might be an encouraging potential source of prognostic markers. Our study revealed that immune cells with tumor-killing effect in high-risk group constituted smaller proportion than their counterparts in low-risk group, which is consistent with a previous research on gastric cancer[52]. Our results revealed the distribution differences of tumor-infiltrating immune cells in TP53-mutated MM patients according to the risk. The association between the immune microenvironment and risk stratification in MM patients should be investigated in future study.

There existed several unavoidable limitations in our study. Firstly, all enrolled patients were derived only from the GEO database with limited sample size. We could not retrieve suitable cohorts from other database, such as the Cancer Genome Atlas (TCGA) database. Secondly, like most of the public databases, the GEO database lacks the important laboratory results about individual MM patients, such as hemoglobin, creatinine, free light chain,
monoclonal immunoglobulin and many other prognostic parameters. Thirdly, information with regard to disease progression, relapse or recurrence, infection, comorbidities and complications were neither recorded nor collected in the GEO database. What’s more, the specific therapeutic regimens, including drug doses and administration frequency were also not recorded. Fourthly, the DEARG risk model was established on a retrospective cohort and we were unable to find external validation datasets to further verify the accuracy and robustness of our model, which needs further validation in other independent prospective cohorts. Lastly, functional experiments both in vivo and in vitro are warranted in the future to explore the mechanisms of DEARGs in our model.

In conclusion, our study identified multiple ARGs which were correlated with the survival outcomes of TP53-mutated patients. By incorporating \textit{CASP8}, \textit{MAPK8} and \textit{RB1CC1}, we established a three-ARGs risk signature which turned out to be an independent prognostic factor. What’s more, by integrating the risk model with ISS stage, we constructed a nomogram which was testified to perform well in predicting long-term survival of MM patients with TP53 mutations.

**Data Availability**

The data that support the results of our study are available in Gene Expression Omnibus (GEO) datasets at https://www.ncbi.nlm.nih.gov/gds/. All original data throughout our manuscript are available upon reasonable request by communicating with corresponding author.
Conflict of Interest

All authors declare no conflict of interest.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (81970190, 81900207), the Innovative Chain (Group) in Key Industry of Shaan’xi Province of China (2019ZDLSF02-02) and the Translation Research Grant of NCRCH (2020ZKMC01).

References


Figure Legends

Figure 1: DEGs between TP53-mutated patients and TP53-unmutated patients. (A) Volcano plot of DEGs. Red nodes suggest up-regulated genes. Green nodes suggest down-regulated genes. Black nodes suggest genes without differences. (B) Heatmap of top 100 DEGs. (C) The Venn diagram of 3329 DEGs (Blue area) and 222 ARGs (Orange area). ARG, autophagy-related gene; DEG, differentially expressed gene; HAD, human autophagy database.

Figure 2: Functional enrichment analysis of 51 DEARGs. (A) KEGG analysis reveals the signaling pathway in which DEARGs get involved. (B) GO analysis reveals the biological process in which DEARGs get involved. DEARG, differentially expressed autophagy-related gene; GO, gene ontology; KEGG, Kyoto encyclopedia of genes and genomes.

Figure 3: The Kaplan-Meier curve for the expression of CASP8, MAPK1, RB1CC1 in TP53-mutated multiple myeloma patients.

Figure 4: Prognostic analysis of three-DEARGs risk-score model in the training set. (A) Time-dependent ROC curve analyses indicate AUC values for 5-year (red), 8-year (yellow) and 10-year (green) OS. (B) Kaplan-Meier OS curve of the high-risk (red) and low-risk (green) group. (C) The risk-score distribution curve of high-risk (red) and low-risk (green) TP53-mutated patients in the OS model. (D) Scattered plot reveals the OS status of TP53-mutated patients. Red dots suggest dead patients while green dots suggest patients who remain alive. (E) Heatmap of the expression profiles of three DEARGs in the high- and low-risk group. AUC, area under the curve; DEARG, differentially expressed autophagy-related gene; OS, overall survival; ROC,
receiver-operating characteristic curve

Figure 5: The nomogram and calibration plot for predicting the OS. (A) The nomogram integrating the risk score and ISS-stage. (B) The calibration plot to testify the consistency of 8-year and 10-year OS between the predicted outcome and actual results. ISS, international staging system; OS, overall survival

Figure 6: Immune cell infiltration landscapes in high- and low-risk MM patients with TP-53 mutations according to DEARGs. (A) Unsupervised clustering of 22 immune-infiltrating cells in 467 patients in the high and low-risk group. (B) Differences of immune cell infiltration abundances between high- and low-risk patients. ns denotes not significant (P>0.05), *denotes p<0.05, **denotes p<0.01, ***denotes p<0.001, ****denotes p<0.0001. DEARG, differentially expressed autophagy-related gene; MM, multiple myeloma.
Table 1. Characteristics of 557 TP53-mutated patients

<table>
<thead>
<tr>
<th></th>
<th>Number(Percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;65</td>
<td>338(60.68%)</td>
</tr>
<tr>
<td>≥65</td>
<td>129(23.16%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>90(16.16%)</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>291(52.24%)</td>
</tr>
<tr>
<td>Female</td>
<td>176(31.60%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>90(16.16%)</td>
</tr>
<tr>
<td><strong>ISS Stage</strong></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>69(12.39%)</td>
</tr>
<tr>
<td>II</td>
<td>305(54.76%)</td>
</tr>
<tr>
<td>III</td>
<td>84(15.08%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>99(17.77%)</td>
</tr>
</tbody>
</table>

ISS: International staging system
Table 2. Univariate and multivariate Cox regression analyses of OS

<table>
<thead>
<tr>
<th></th>
<th>Univariate analysis</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR</td>
<td>95%CI</td>
<td>P-value</td>
</tr>
<tr>
<td>Age</td>
<td>1</td>
<td>0.99-1.02</td>
<td>0.5</td>
</tr>
<tr>
<td>Gender</td>
<td>0.95</td>
<td>0.76-1.20</td>
<td>0.68</td>
</tr>
<tr>
<td>ISS Stage</td>
<td>1.9</td>
<td>1.57-2.30</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Risk Score(High vs Low)</td>
<td>3.29</td>
<td>2.35-4.60</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

ISS: International staging system; HR: Hazard ratio; CI: Confidence interval
<table>
<thead>
<tr>
<th>Cell type</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasma cells</td>
<td>4.98E-17*</td>
</tr>
<tr>
<td>Monocytes</td>
<td>1.41E-16*</td>
</tr>
<tr>
<td>Resting mast cells</td>
<td>1.03E-15*</td>
</tr>
<tr>
<td>Activated NK cells</td>
<td>9.36E-10*</td>
</tr>
<tr>
<td>Activated dendritic cells</td>
<td>3.72E-09*</td>
</tr>
<tr>
<td>Resting NK cells</td>
<td>6.51E-09*</td>
</tr>
<tr>
<td>Memory B cells</td>
<td>1.00E-08*</td>
</tr>
<tr>
<td>Activated mast cells</td>
<td>9.34E-06*</td>
</tr>
<tr>
<td>Naïve CD4+ T cells</td>
<td>1.57E-03*</td>
</tr>
<tr>
<td>Follicular helper T cells</td>
<td>3.86E-03*</td>
</tr>
<tr>
<td>Gamma delta T cells</td>
<td>8.32E-03*</td>
</tr>
<tr>
<td>Macrophages-M0</td>
<td>1.99E-02*</td>
</tr>
<tr>
<td>CD8+ T cells</td>
<td>2.41E-02*</td>
</tr>
<tr>
<td>Resting dendritic cells</td>
<td>1.63E-01</td>
</tr>
<tr>
<td>Naïve B cells</td>
<td>1.70E-01</td>
</tr>
<tr>
<td>Eosinophils</td>
<td>2.35E-01</td>
</tr>
<tr>
<td>Neutrophils</td>
<td>3.49E-01</td>
</tr>
<tr>
<td>Regulatory T cells(Tregs)</td>
<td>3.96E-01</td>
</tr>
<tr>
<td>Macrophages-M1</td>
<td>4.96E-01</td>
</tr>
<tr>
<td>Macrophages-M2</td>
<td>5.46E-01</td>
</tr>
<tr>
<td>Activated memory CD4+ T cells</td>
<td>5.62E-01</td>
</tr>
<tr>
<td>Resting memory CD4+ T cells</td>
<td>7.12E-01</td>
</tr>
</tbody>
</table>

*denotes P<0.05 with statistical significance by T-test
Figure 1: DEGs between TP53-mutated patients and TP53-unmutated patients. (A) Volcano plot of DEGs. Red nodes suggest up-regulated genes. Green nodes suggest down-regulated genes. Black nodes suggest genes without differences. (B) Heatmap of top 100 DEGs. (C) The Venn diagram of 3329 DEGs (Blue area) and 222 ARGs (Orange area).

ARG, autophagy-related gene; DEG, differentially expressed gene; HAD, human autophagy database
Figure 2: Functional enrichment analysis of 51 DEARGs. (A) KEGG analysis reveals the signaling pathway in which DEARGs get involved. (B) GO analysis reveals the biological process in which DEARGs get involved. DEARG, differentially expressed autophagy-related gene; GO, gene ontology; KEGG, Kyoto encyclopedia of genes and genomes.
Figure 3: The Kaplan-Meier curve for the expression of CASP8, MAPK1, RB1CC1 in TP53-mutated multiple myeloma patients.
Figure 4: Prognostic analysis of three-DEARGs risk-score model in the training set. (A) Time-dependent ROC curve analyses indicate AUC values for 5-year (red), 8-year (yellow), and 10-year (green) OS. (B) Kaplan-Meier OS curve of the high-risk (red) and low-risk (green) group. (C) The risk-score distribution curve of high-risk (red) and low-risk (green) TP53-mutated patients in the OS model. (D) Scattered plot reveals the OS status of TP53-mutated patients. Red dots suggest dead patients while green dots suggest patients who remain alive. (E) Heatmap of the expression profiles of three DEARGs in the high- and low-risk group. AUC, area under the curve; DEARG, differentially expressed autophagy-related gene; OS, overall survival; ROC, receiver-operating characteristic curve.
Figure 5: The nomogram and calibration plot for predicting the OS. (A) The nomogram integrating the risk score and ISS-stage. (B) The calibration plot to testify the consistency of 8-year and 10-year OS between the predicted outcome and actual results. ISS, international staging system; OS, overall survival.
Figure 6: Immune cell infiltration landscapes in high- and low-risk MM patients with TP-53 mutations according to DEARGs. (A) Unsupervised clustering of 22 immune-infiltrating cells in 467 patients in the high and low-risk group. (B) Differences of immune cell infiltration abundances between high- and low-risk patients. ns denotes not significant (P>0.05), *denotes p<0.05, **denotes p<0.01, ***denotes p<0.001, ****denotes p<0.0001. DEARG, differentially expressed autophagy-related gene; MM, multiple myeloma.