Original Article

Integrating Lysosomal Genes and Immune Infiltration for Multiple Myeloma Subtyping and Prognostic Stratification

(multiple myeloma / immune score / lysosome / prognostic model / autophagy)

SHU DENG, JINGJING XIANGANG, ZHIYIN ZHENG, JIANPING SHEN

Department of Hematology, The First Affiliated Hospital of Zhejiang Chinese Medical University (Zhejiang Provincial Hospital of Chinese Medicine), Hangzhou 310006, China

Abstract. Lysosomes are crucial in the tumour immune microenvironment, which is essential for the survival and homeostasis in multiple myeloma (MM). Here, we aimed to identify lysosome-related genes for the prognosis of MM and predicted their regulatory mechanisms. Gene expression profiles of MM from the GSE2658 and GSE57317 datasets were analysed. Lysosome-related differentially expressed genes (DEGs) were identified and used for molecular subtyping of MM patients. A prognostic model was constructed using univariate Cox regression and LASSO regression analyses. The relationship between prognostic genes, immune cell types, and autophagy pathways was assessed through correlation analysis. RT-qPCR was performed to validate the expression of prognostic genes in MM cells. A total of 9,954 DEGs were identified between high and low immune score groups, with 213 intersecting with lysosomal genes. Molecular subtyping revealed two distinct MM subtypes with significant differences in immune cell types and autophagy pathway activities. Five lysosome-related DEGs (*CORO1A*, *ELANE*, *PSAP*, *RNASE2*, and *SNAPIN*) were identified as significant prognostic markers. The prognostic model showed moderate predictive accuracy with AUC values up to 0.723. Prognostic genes demonstrated significant correlations with various immune cell types and autophagy pathways. Additionally, *CORO1A*, *PSAP* and *RNASE2* expression was up-regulated in MM cells, while *ELANE* and *SNAPIN* were down-regulated. Five lysosomal genes in MM were identified, and a new risk model for prognosis was developed using these genes. This research could lead to discovering important gene markers for the treatment and prognosis of MM.

Introduction

Multiple myeloma (MM) is a malignant transformation of plasma cells that synthesize and secrete immunoglobulin (Mahindra et al., 2010). Globally, MM is diagnosed in an estimated 588,162 people every year and accounts for approximately 2 % of all cancer deaths (Padala et al., 2021; Cowan et al., 2022). A statistic based on 169 cases reported that the median survival of patients with MM was 33 months and the five-year overall survival was 15.5 % (Acquah et al., 2019). Although various therapy strategies have been implemented in the clinical treatment of MM, the overall survival (OS) of MM patients is still not ideal (Fonseca et al., 2017; Moreau et al., 2021). How to improve the survival rate of MM patients and explore the biological mechanism affecting the prognosis is an urgent problem to be solved.

Lysosomes are single-membrane organelles with acid hydrolases that break down cellular materials and help maintain the cell balance (Mahapatra et al., 2021). These organelles, through their involvement in amino acid homeostasis, survival signalling and immune evasion, substantially contribute to tumour progression (Tang et al., 2020). Lysosomes play a critical role in autophagy, degrading nonessential macromolecules to support the increased nutrient and energy demands of cancer cells (Amaravadi et al., 2019). Under conditions like ischaemia and hypoxia, they activate autophagy to counteract nutrient shortages, recycling materials with their hydrolases (Romao et al., 2013; Pu et al., 2016). In addition, lysosomes govern the transport and degradation of cell surface molecules, including cytotoxic glycoprotein T lymphocyte antigen 4 (CTLA-4), thereby playing a

Received January 25, 2024. Accepted April 25, 2024.

This study was supported by a Chinese Medicine Research Fund Project of Zhejiang Province (number 2021ZB092) and the Zhejiang Province Natural Science Foundation (number LQ18H080003).

Corresponding author: Jianping Shen, Department of Hematology, The First Affiliated Hospital of Zhejiang Chinese Medical University (Zhejiang Provincial Hospital of Chinese Medicine), No. 54 Youdian Road, Hangzhou 310012, China. E-mail: Shenjianping2006@163.com

Abbreviations: AUC – area under the curve, DEGs – differentially expressed genes, GEO – Gene Expression Omnibus, HGNC – HUGO Gene Nomenclature Committee, MM – multiple myeloma, OS – overall survival, PBMCs – peripheral blood mononuclear cells.

pivotal role in the immune regulation of cancers (van Kasteren and Overkleeft, 2014; Zhu et al., 2020b). As research in tumour metabolism and gene therapy advances, lysosomal genes are increasingly recognized as highly valuable targets for cancer therapy (Huang et al., 2023). Nonetheless, the correlation between lysosomes and MM is not yet fully understood.

In this study, based on the data in the Gene Expression Omnibus (GEO) database, the immune and stromal scores were evaluated in MM patients, followed by exploration of differentially expressed genes (DEGs) between high and low score groups. Lysosomal genes were obtained by merging data from the hLGDB and MSigDB databases. Then, investigation of prognostic genes, enrichment analysis and immune analysis were performed, followed by establishment of a prognostic model. Finally, the associations of the prognostic model with MM patient survival were investigated. We hoped to reveal the potential molecular mechanism associated with MM prognosis and investigate a promising prognostic model of MM based on lysosome-related DEGs.

Material and Methods

Microarray data and pre-processing

The microarray datasets GSE2658 (Hanamura et al., 2006), GSE57317 (Heuck et al., 2014) and GSE4581 were acquired from the GEO database using the GEOquery package (Davis and Meltzer, 2007) in R software based on the platform of GPL570 Affymetrix Human Genome U133 Plus 2.0 Array. GSE2658 and GSE57317 served as training datasets; GSE4581 served as an external validation dataset. The pre-processing of two datasets (GSE2658 and GSE57317) was performed using the hgu133plus2.db (Carlson et al., 2016) package in R, followed by reannotation based on the annotation information provided by the HUGO Gene Nomenclature Committee (HGNC) database (Gray et al., 2016). We employed the 'sva' package in R to integrate two gene chip datasets and mitigate batch effects, as part of preliminary research encompassing a total of 614 MM samples.

Assessment of immune scores

Through the 'estimate' package in R (Yoshihara et al., 2013), this method aligns and filters genes in the expression profile against a set of common genes previously identified by researchers. Following this, the Affymetrix platform is employed to calculate immune scores. Based on these scores, samples are classified: those surpassing the lower quartile are designated as high immune samples (IS_high), and the remaining as low immune samples (IS_low).

Identification of lysosome-related DEGs in MM

Employing the 'limma' package in R, this study identified DEGs between IS_high and IS_low groups. The selection criteria for these DEGs were a P value < 0.05 and an absolute log2 fold change ($|\log2FC| > 0.1$. Subsequently, the 'ggplot2' package, also in R, was utilized to generate both volcano plots and clustering heatmaps for these DEGs. Additionally, lysosomal genes were consolidated from the hLGDB and MSigDB databases. This consolidated gene set was then intersected with the identified DEGs using the R's 'base' package, and Venn diagrams were created to visualize these intersections.

Molecular subtyping and immune infiltration analysis

The 'ConsensusClusterPlus' package in R (Yu et al., 2012) was employed for consensus clustering analysis based on the expression of intersected genes within the MM gene chip datasets. The objective was to elucidate the relationship between these intersected genes and the various subtypes of MM. The clustering analysis was performed with the number of clusters (k) varying from 2 to 10, and the optimal clustering configuration was determined using the cumulative distribution function (CDF). Subsequently, the study utilized the ssgsea method (Wilkerson and Hayes, 2010) to conduct cell infiltration analysis in both immune score groups (IS_high vs IS_low) and across different MM subtypes. Statistical differences between groups were assessed using the *t*-test to calculate P values, focusing on identifying cells that exhibited significant differences in both scoring and subtype groups ($P < 0.05$).

Gene set variation analysis (GSVA)

GSVA is capable of transforming a gene-sample data matrix into a gene set-sample matrix. This transformation facilitates further analysis of the enrichment of gene sets in each sample. For this study, 26 autophagy-related pathways were downloaded from the MSigDB database (Liberzon et al., 2015) to serve as the background gene sets. Using the GSVA package in R, enrichment analysis was performed to obtain activity scores of these pathways in the samples. Subsequently, the *t*-test method was used to examine the differences in these 26 pathways across different MM subtypes ($P < 0.05$).

Univariate Cox regression analysis

Univariate Cox regression analysis was conducted on intersected genes using the 'survival' package in R. The analysis focused on patients with MM, utilizing their gene expression profiles and survival data. The aim was to explore the relationship between genes intersected in the lysosomal pathway and the prognosis of MM. Utilizing the log-rank test, genes that demonstrated a significant correlation with the prognosis of MM were identified, each surpassing the significance threshold of $P < 0.05$. A total of 53 genes were screened.

Construction and validation of the prognostic model

To construct a robust prognostic risk model, a least absolute shrinkage and selector operation (LASSO) analysis was conducted to filter out key prognostic genes through the glmmet package in R. The lambda value was screened through cross-validation, and the model was constructed using the lambda.min value. The gene expression matrix was extracted from this model, and each sample's risk score was calculated by the following formula:

$$\text{Risk score} = \sum_{j=1}^{n} expj \times coefj$$

where exp represents the expression of the corresponding genes, coef is the regression coefficient, j stands for gene.

The dataset containing gene expression and survival data from MM patients was randomly divided equally into a training set and an internal validation set. The training set was analysed using LASSO regression via the Cox method in the R's 'glmnet' package, resulting in the development of a prognostic model based on identified prognostic genes. The validation set and the external validation set (GSE4581) were utilized to validate the effectiveness and accuracy of the newly established prognostic model. The prognostic risk was assessed using the prognostic model, leading to the stratification of all patients into high- and low-risk groups based on the median value of the prognostic risk score. Subsequently, the Kaplan-Meier (KM) method was employed for survival analysis of these groups, utilizing the 'survival' and 'survminer' packages in R (Rizvi et al., 2019; Wang et al., 2020). Additionally, time-dependent receiver operating characteristic (ROC) curve analyses were performed to investigate the survival of patients in 1-, 2-, and 3-year intervals using the 'survivalroc' package (Zhu et al., 2020a).

Correlation analysis

Pearson correlation analysis, a commonly employed statistical technique, assesses the strength and direction of a linear relationship between two variables. This method hinges on the principle of covariance. It involves computing the correlation coefficient, which ranges from −1 to 1, by dividing the covariance of the variables by the product of their respective standard deviations. For this study, the Pearson correlation was calculated to explore the relationships between prognostic genes and their correlation with immune cells as well as pathways related to autophagy, maintaining a significance level of $P < 0.05$.

Cell culture

Peripheral blood mononuclear cells (PBMCs) from healthy individuals (QCLL-1012114, Hefei Wanwu Biotechnology Co., Ltd., Anhui, China) and multiple myeloma cell lines MM1S (CTCC-004-0048, Zhejiang Meisen Cell Technology Co., Ltd., Zhejiang, China), RPMI-8226 (CTCC-400-0191, Zhejiang Meisen Cell Technology) and U266 (CTCC-400-0175, Zhejiang Meisen Cell Technology) were cultured. The cells were grown in RPMI 1640 medium supplemented with 10 % foetal bovine serum (10099-141, Gibco-Thermo Fisher Scientific, Waltham, MA), 100 U/ml penicillin, and 100 μg/ml streptomycin. The culture was maintained at 37 °C in a humidified incubator with 5 % CO_2. The medium was changed every 2–3 days. When cell density reached approximately 1.5×10^6 cells/ml, the cells were subcultured.

RT-qPCR

Intracellular RNA was extracted by Trizol reagent, and the quality of the extracted RNA was detected by Nano-Drop 8000 (Thermo Fisher Scientific, Waltham, MA). The extracted RNA was reverse transcribed to cDNA according to manufacturer's instructions to detect the expression of following targets. The SYBR Green PCR Master Mix (Lifeint, Xiamen, China) and the CFX96 Touch Real-Time PCR Detection System (Bio-Rad, Hercules, CA) were utilized to perform RT-qPCR reactions. Primer sequences are shown in Table S1. The CT values of each gene were counted, and the relative expression of prognostic genes was analysed according to the 2-ΔCt method using *GAPDH* as the internal reference gene.

Statistical analysis

The experiment was replicated more than three times to ensure reliability. Data are expressed as the mean ± standard deviation. To assess intergroup differences, the *t*-test and univariate analysis of variance (ANOVA) were employed. Pearson correlation was used to explore the relationships. Statistical analyses were conducted using the R software (version 4.3.2, R Foundation for Statistical Computing, Vienna, Austria) and GraphPad 7.0 software, with a P value threshold of 0.05 set for determining statistical significance.

Results

Identification of lysosome-related DEGs in MM

The gene expression profiles of MM were extracted from the GSE2658 and GSE57317 datasets. Utilizing the 'estimate' package within R version 4.3.2, gene alignment was performed on expression profile files. This procedure entailed a meticulous filtering phase, succeeded by the selection of the Affymetrix platform to calculate immune scores. Subsequently, samples surpassing the lower quartile in immune scores were categorized as high-immune samples (IS_high), with the remaining classified as low-immune samples (IS_low). A total of 9,954 DEGs were identified between IS_high and IS_low groups. Among these, 6,532 genes were significantly up-regulated, while 3,422 genes exhibited

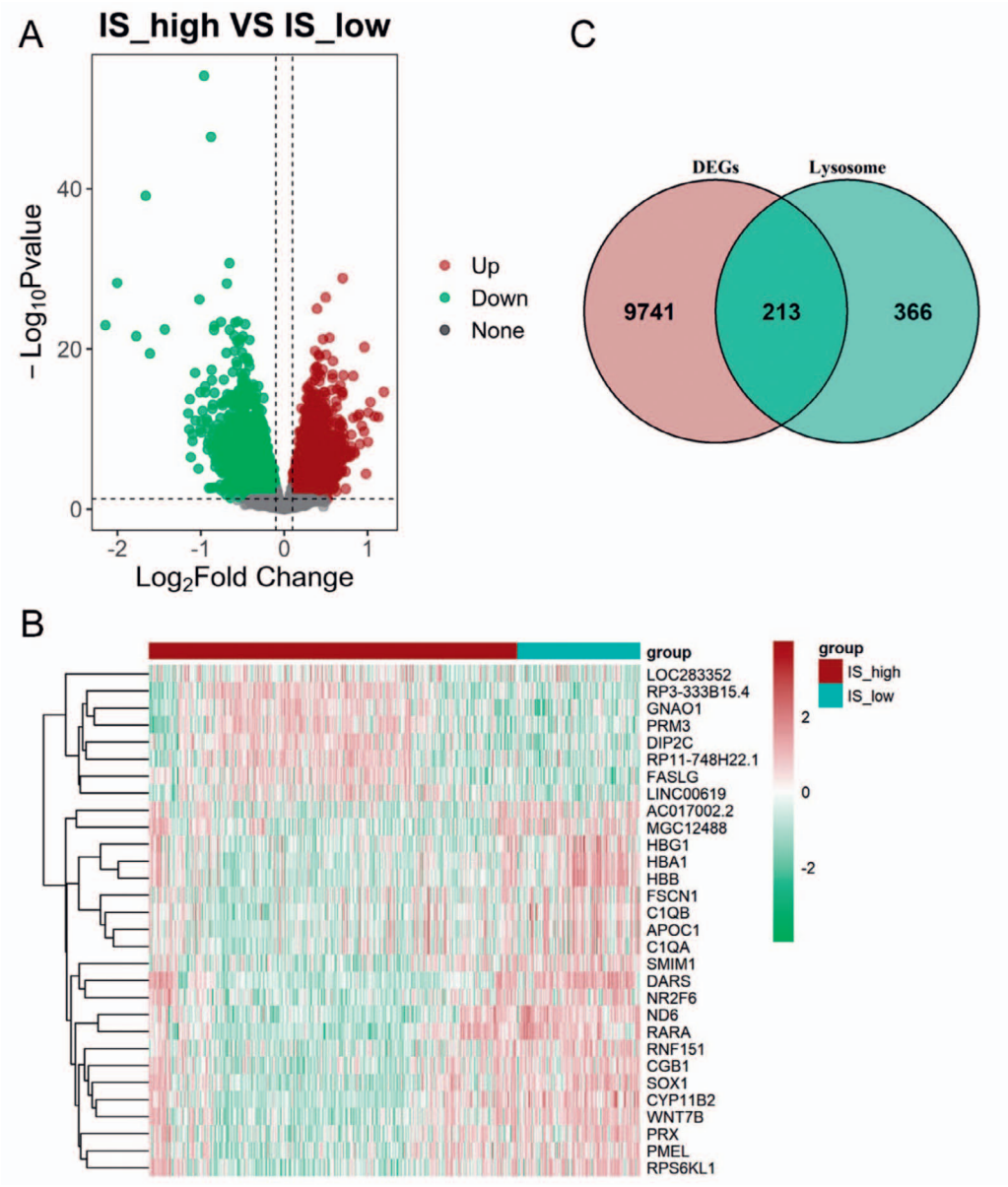

Fig. 1. Identification of lysosome-related DEGs in MM. **(A)** Volcano plots of the distribution of DEGs in high-immune (IS_high) and low-immune (IS_low) sample groups in MM. **(B)** Heatmap of top 30 DEGs in IS_high and IS_low sample groups in MM. **(C)** Intersection of DEGs and lysosomal genes in MM.

significant down-regulation (Fig. 1A). The top 30 genes with the most significant variabilities for each dataset between IS_high and IS_low groups were clustered in the heatmap separately (Fig. 1B). A total of 579 lysosomal genes were obtained by merging data from the hLGDB and MSigDB databases. The intersection of DEGs and lysosomal genes yielded a total of 213 intersecting genes (Fig. 1C).

Molecular subtyping of MM via lysosome-related DEGs

It could be seen that the best clustering results were achieved when 256 MM patients were clustered into two molecular subtypes (k = 2) based on the expression patterns of 213 lysosome-related DEGs (Fig. 2A). The PCA results showed the reasonableness of the clustering result, exhibiting a good internal consistency and stability (Fig. 2B). The result of K-M curve revealed that the prognosis was better for subtype 2 than subtype 1 (Fig. 2C). Subsequently, using the ssGSEA method and *t*-test, we analysed the differences in 23 types of immune cells across various subtypes and between IS_high and IS_low groups. Ultimately, 12 immune cell types were identified that exhibited significant differences both between different subtypes and within the IS_high and IS_low groups ($P < 0.01$; Fig. 2D–E). Additionally, the GSVA analysis of 26 autophagy pathways revealed significant

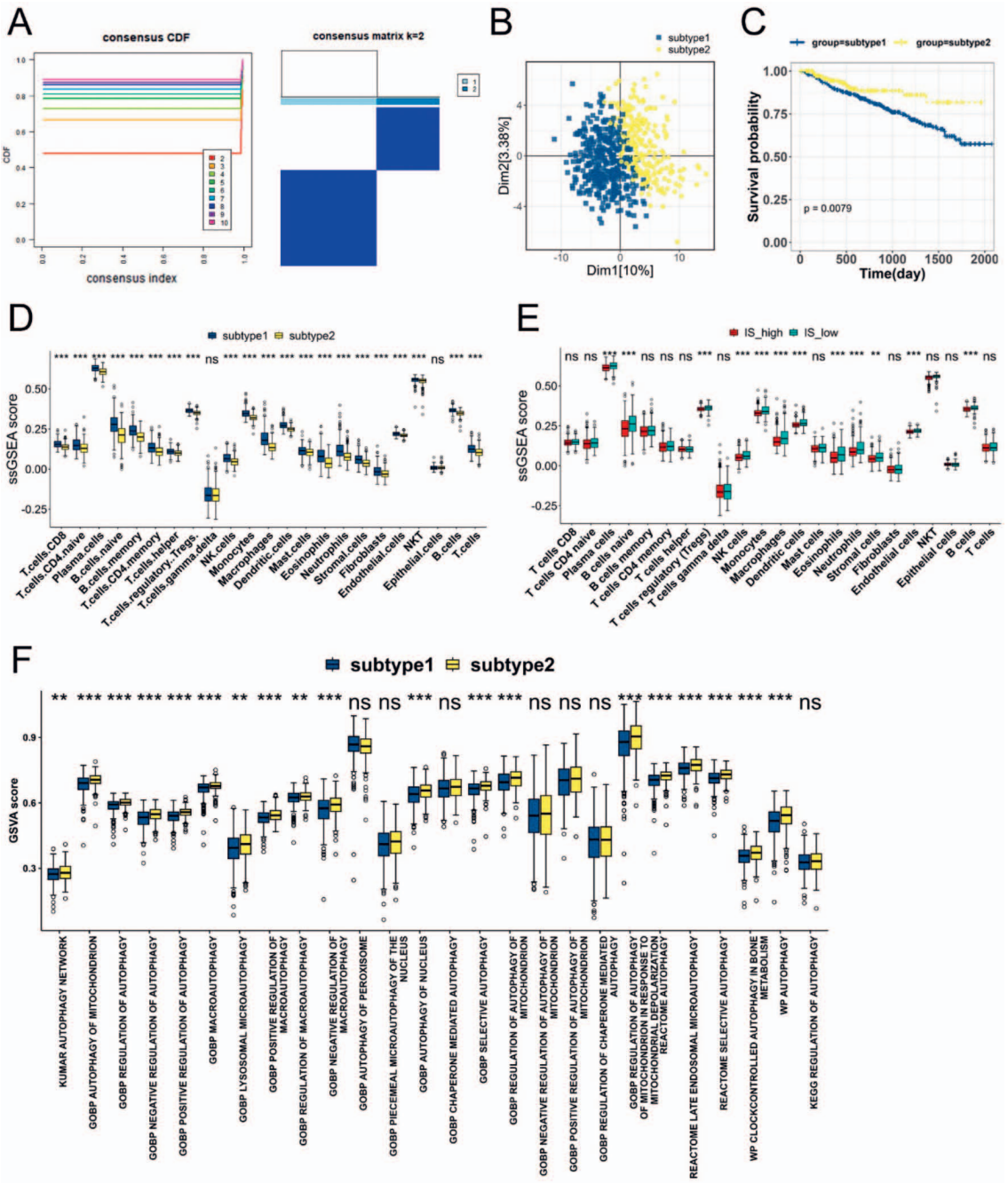

Fig. 2. Molecular subtyping of MM via lysosome-related DEGs. **(A)** Molecular subtyping of MM using lysosome-related DEGs. **(B)** Principal component analysis (PCA) of two MM subtypes. **(C)** Kaplan-Meier survival curves for two MM subtypes. **(D)** Immune cell type differences in two MM subtypes. **(E)** Immune cell type differences within IS_high and IS_low groups. **(F)** GSVA analysis of autophagy pathways in MM subtypes.

score differences across subtypes in 19 pathways, indicating distinct subtype-specific autophagic profiles ($P < 0.01$; Fig. 2F).

Establishment of the prognostic model

We conducted a univariate Cox regression analysis based on the expression of the 213 lysosome-related DEGs using the 'survival' package in R version 4.3.2. Through log-rank testing, we successfully identified 53 genes significantly associated with the prognosis of multiple myeloma ($P < 0.05$; Fig. 3A). Using lambda.min through cross-validation in LASSO, a total of five lysosome-related DEGs were screened out as prognostic markers, which were *CORO1A*, *ELANE*, *PSAP*, *RNASE2*, and *SNAPIN* (Fig. 3B). Based on the five characteristic genes, we constructed a prognostic risk model for MM patients. Risk score = 0.409861719**CORO1A*-0.187476224**ELANE*-0.782219869**PSAP*-

0.064489938**RNASE2**+0.251009196**SNAPIN*. Based on the median risk, the MM patients were divided into the high- and low-risk groups. Prognostic gene analysis involved randomly dividing MM gene expression and survival data into two equal parts. One half was used as the training set, and the other as the internal validation set, to validate the developed prognostic model. The result of K-M curve revealed that the prognosis was better for lower risks in both training and internal validation sets (Fig. 3C). The ROC analysis revealed that our prognostic model exhibited an area under the curve (AUC) of 0.665, 0.679, and 0.723 for 1-, 2-, and 3-year intervals, respectively, in the training set (Fig. 3D). In the internal validation set, the AUCs were 0.566, 0.606, and 0.633 at the corresponding time points (Fig. 3D). To further validate the ability of the prognostic model, an external validation set was also used to evaluate the K-M curve and ROC curve. The results also showed higher survival in low-risk groups, with AUC of 0.609, 0.656, and 0.691 for 1-, 2-, and 3-year intervals, respectively (Fig. 3C–D). These results indicate that the prognostic risk model demonstrates moderate predictive accuracy.

Correlation analysis of prognostic genes and immune microenvironment

Correlation analysis among 12 immune cell types, exhibiting significant differences across subtypes and IS_high/low groups, and five prognostic genes revealed notable associations. The analysis indicated that three genes – *ELANE*, *PSAP* and *RNASE2* – were significantly positively correlated with all 12 immune cell types ($P < 0.05$; Fig. 4A–C). *SNAPIN* showed a significant positive correlation with Tregs, monocytes and endothelial cells ($P < 0.05$; Fig. 4D). *CORO1A* was found to have a significant positive correlation with six cell types: Tregs, monocytes, endothelial cells, dendritic cells (DC), B cells naive, and B cells ($P < 0.05$; Fig. 4E). Additionally, Fig. S1 presents the significant correlation between 19 autophagy pathways, which exhibit significant differ-

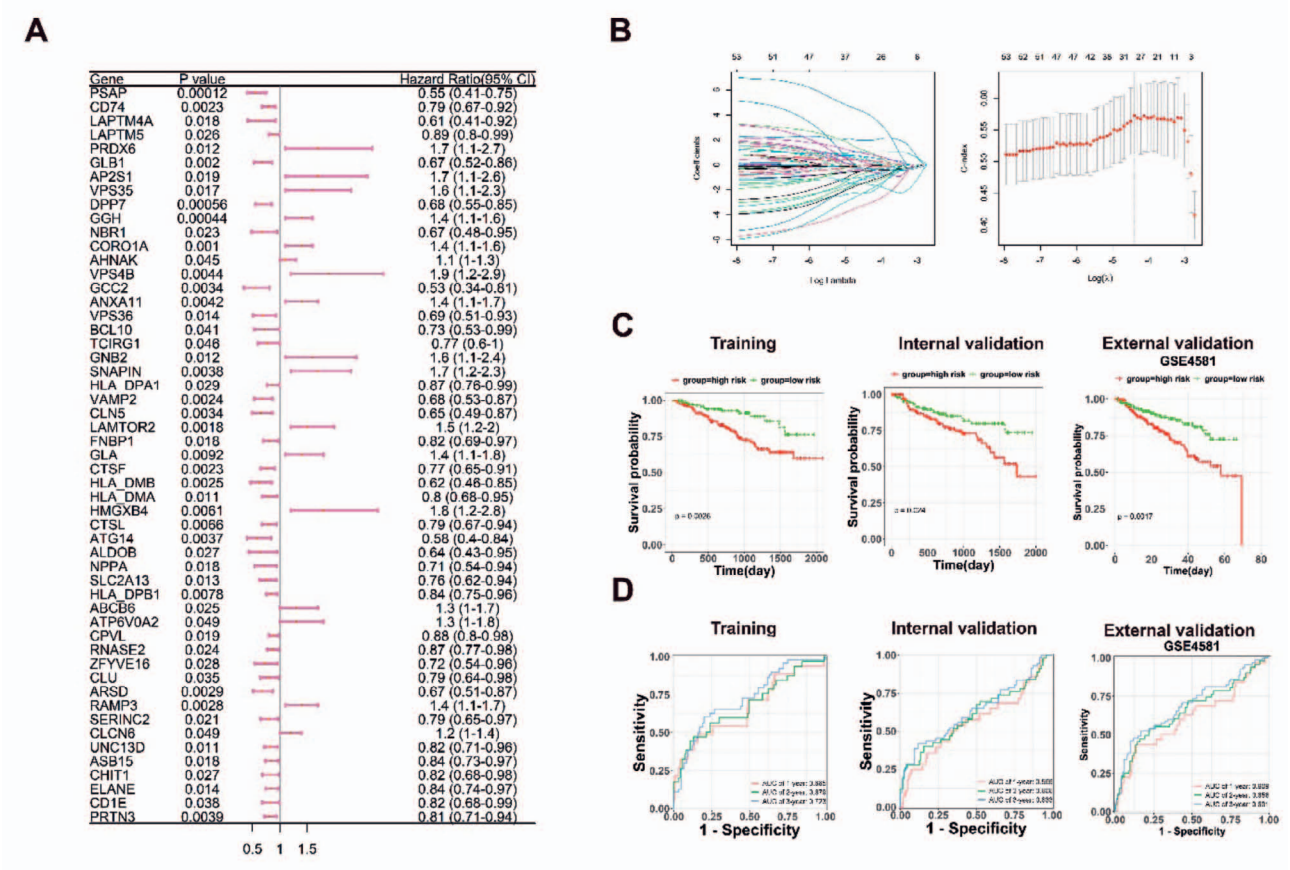

Fig. 3. Establishment of the prognostic model. **(A)** Forest plot of the effect of lysosome-related DEG expression levels on the patient survival by Cox regression analysis. **(B)** LASSO regression analysis of MM prognostic markers. Left, correlation between the LASSO regression coefficient and lambda. Right, tenfold cross-validation for turning parameter selection in the LASSO Cox regression model. Lambda (λ) is the turning parameter. The partial likelihood deviance is plotted in log(λ), in which vertical lines are shown at the optimal values by minimum criteria and 1 − SE criteria. **(C)** Kaplan-Meier survival curves for MM risk groups. **(D)** ROC analysis of the MM prognostic model.

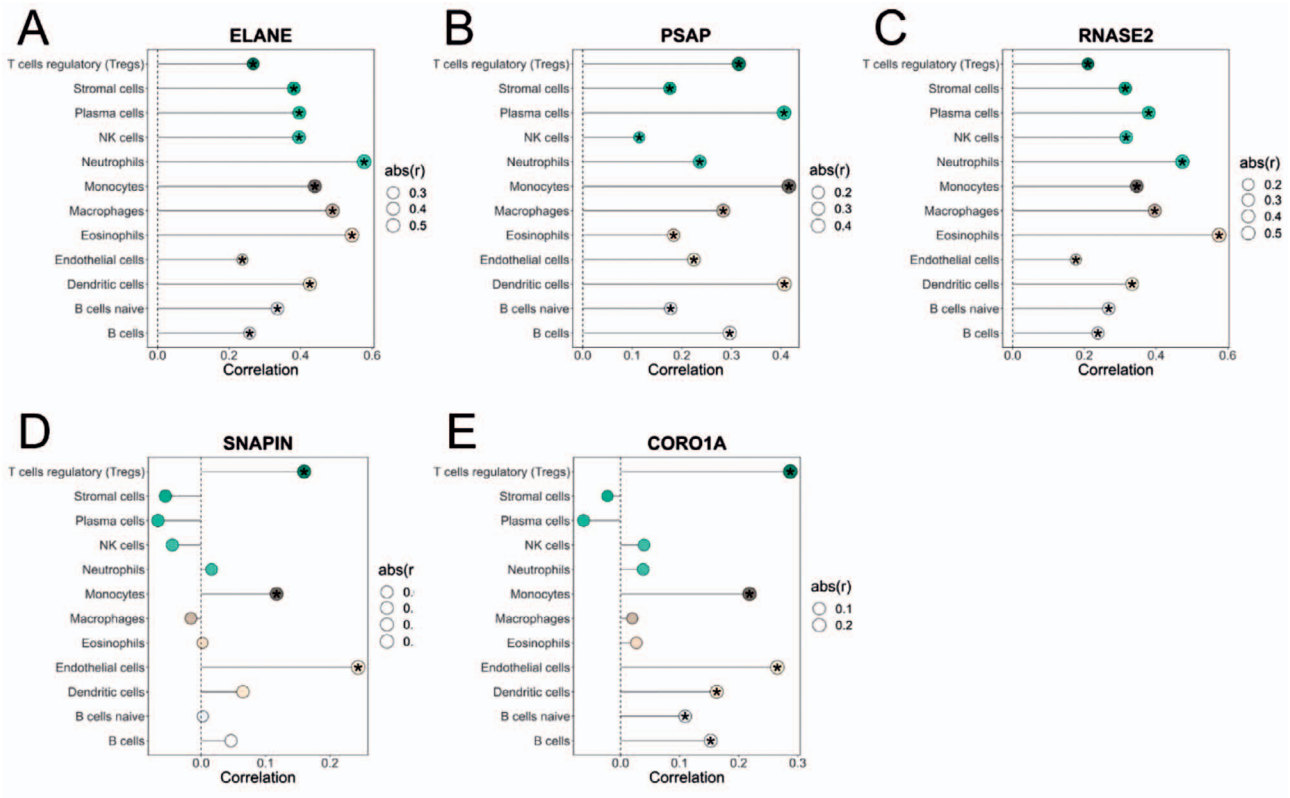

Fig. 4. Correlation between prognostic genes and immune cell types in MM. These plots illustrate the significant correlations between the five identified prognostic genes (*ELANE* **(A)**, *PSAP* **(B)**, *RNASE2* **(C)**, *SNAPIN* **(D)**, *CORO1A* **(E)**) and 12 immune cell types, revealing intricate interactions within the immune microenvironment of MM.

ences across various subtypes, and five prognostic genes ($P < 0.05$; Fig. S1).

Expression and validation of prognostic genes

The analysis of data from the GEO dataset GSE125361 indicated that the expression levels of *PSAP* were markedly elevated in MM samples relative to normal tissue samples ($P < 0.01$; Fig. 5A). Conversely, *RNASE2*, *ELANE* and *SNAPIN* exhibited reduced expression in MM samples compared to their normal counterparts ($P < 0.05$; Fig. 5A). Consistently, we found that the *CORO1A*, *PSAP* and *RNASE2* expression was much higher in MM cells (MM1S, RPMI-8226, and U266) than in normal cells (PBMCs), while that of *ELANE* and *SNAPIN* was the opposite ($P < 0.05$; Fig. 5B).

Discussion

MM, a prevalent haematologic malignancy, presents intricate outcomes influenced by genetic factors, contributing to the complexity of tumour development and progression (Heider et al., 2021). Although there has been significant advancement in understanding and treating MM (Shah and Mailankody, 2020; Mikkilineni and Kochenderfer, 2021), the pathogenesis of this disease remains elusive. Consequently, exploring the molecular mechanisms underlying MM is both urgent and essential. Prior research has established that inhibiting lysosomal functions can amplify the efficacy of immune checkpoint inhibitors in treating established flank tumours (Yamamoto et al., 2020; Sharma et al., 2022). This opens a novel avenue for investigating lysosomal genes within the tumour immune microenvironment of MM, paving the way for the development of innovative therapeutic targets and mechanisms.

In the present research, a total of 213 lysosome-related DEGs were identified in MM. Based on these DEGs, we defined two MM molecular subtypes and found that patients in subtype 1 had poorer survival than those in subtype 2. An integrative immune infiltration analysis of 23 immune cell types demonstrated distinct infiltration levels of 21 immune cells between subtype 1 and subtype 2. Further, by correlating these findings with different immune cells in the IS_high and IS_low groups of MM, we identified 12 unique immune cell types in MM. This suggests that the composition of the immunological microenvironment may play a significant role in influencing the prognosis of MM. The lysosome, an organelle traditionally regarded as the terminal executor for the degradation of autophagosomal contents, has recently been revealed to also play a pivotal role in the upstream regulation of autophagy (Shen and Mizushima, 2014; Wang et al., 2023). Therefore, we identified 29 autophagy-regulated pathways in subtypes 1 and 2 of MM.

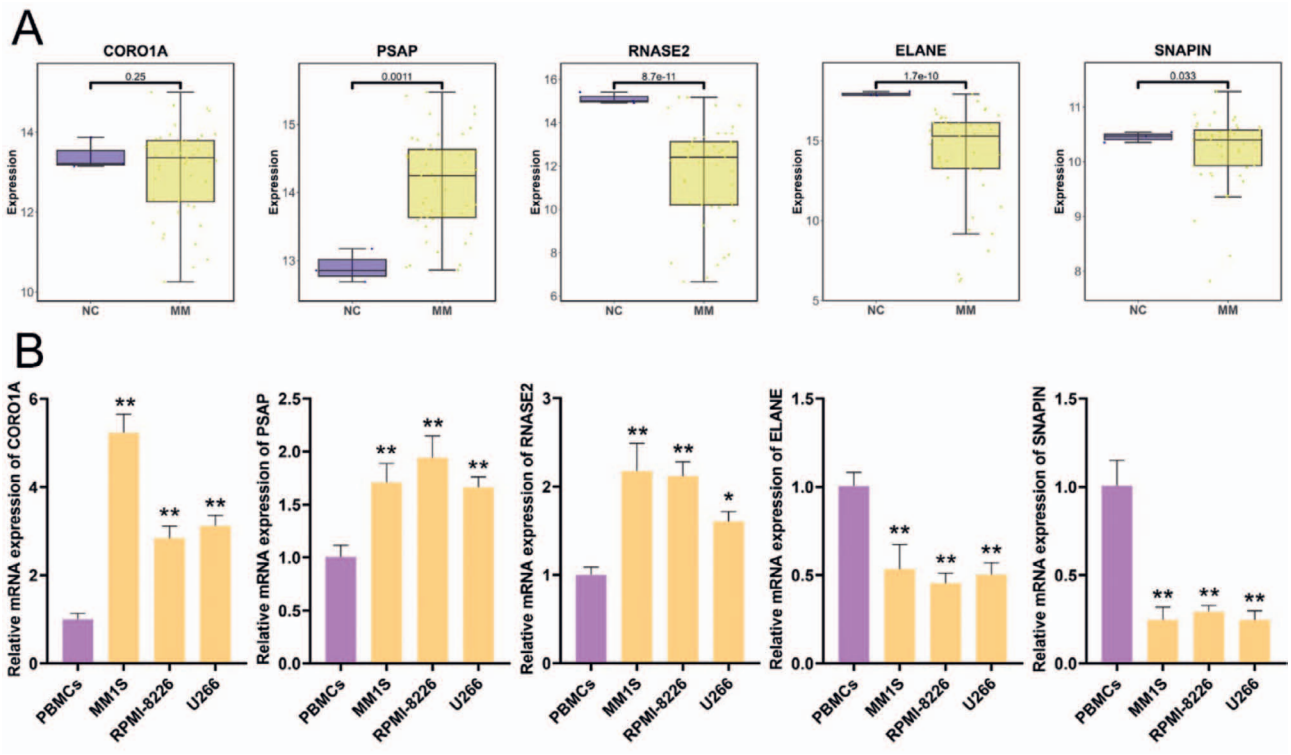

Fig. 5. Expression and validation of prognostic genes. **(A)** Expression levels of prognostic genes in MM and normal tissues. **(B)** Validation of prognostic gene expression in MM cell lines (MM1S, RPMI-8226, U266) compared to normal cells (PBMCs).

Among these, 19 pathways exhibited significant differences between the two subtypes. This differentiation highlights the diversity within the autophagic profile of MM, possibly indicating varied prognostic outcomes and responses to therapy across these subtypes.

To construct a prognostic model, five lysosome-related DEGs were screened out as prognostic markers, namely *ELANE*, *RNASE2*, *CORO1A*, *PSAP*, and *SNAPIN*, the risk score of which had high predictive performance. Compared to a previous study that revealed hub genes in MM (Huang et al., 2021; Pan et al., 2021), we found that *ELANE* and *RNASE2* were also present as hub genes. Huang et al. (2021) found down-regulation of *ELANE* expression in MM, and ROC curve analysis determined that the optimal diagnostic threshold of ELANE was 50 µg/l. Consistently, we also observed lower expression of *ELANE* in MM cell lines. *RNASE2*, a constituent of the RNase A gene family, is primarily involved in immune function and plays a crucial role in activating toll-like receptors (Ostendorf et al., 2020). Pan et al. (2021) found that over-expressed *RNASE2* was associated with worse prognosis in MM patients. For the remaining three genes, *CORO1A* is an autophagosome-lysosome fusion-related gene. Enhanced expression of *CORO1A* in LO2 liver cells was found to impede autophagic flux and increase the levels of inflammatory cytokines following stimulation with palmitic acid (Deng et al., 2024). Our correlation analysis showed that *CORO1A* was significantly associated with 17 autophagy-related pathways. PSAP, a pivotal lysosomal protein, is essential for activating the enzyme glucocerebrosidase and transporting progranulin, thereby significantly influencing lysosomal biogenesis (Tayebi et al., 2020). Prior research has shown that *PSAP* is markedly over-expressed in various cancers such as glioma (Jiang et al., 2019), pancreatic cancer (Miyahara et al., 2022), and colorectal cancer (Robles et al., 2022). Consistent with these findings, our investigation revealed a similar pattern of elevated *PSAP* expression in MM. SNAPIN, which is an adaptor protein in the SNARE core complex, is critical for lysosomal acidification and autophagosome maturation in macrophages. SNAPIN, serving as an adaptor protein within the SNARE core complex, plays a vital role in the acidification of lysosomes and maturation of autophagosomes in macrophages (Shi et al., 2017). Our study demonstrated that *SNAPIN* is significantly correlated with 16 autophagy-related pathways, including lysosomal microautophagy, autophagy of mitochondrion, and autophagy in bone metabolism. In general, the above evidence indicated the role of these key genes in various diseases and lysosomes, whereas the specific role of them in MM, as well as the potential mechanism of them involved in the MM pathogenesis, remain obscure. Hence, more studies focused on these key prognostic genes in MM need to be explored in future.

The study presents certain limitations. Primarily, it is dependent on specific datasets (GSE2658 and GSE57317), potentially limiting the representation of the complete genetic diversity in MM. Additionally, the gene expression analysis, confined to MM cells and tissues, might not comprehensively reflect the intricacies of MM pathogenesis and progression. Furthermore, comprehensive clinical evaluations are required to ascertain the prognostic efficacy of the identified genes.

Conclusion

We identified five lysosome-related DEGs – *ELANE*, *RNASE2*, *CORO1A*, *PSAP*, and *SNAPIN* – as significant prognostic markers for MM. Based on these genes, we developed a prognostic predictive model demonstrating high predictive performance. This model has the potential to improve MM prognostication and inform clinical decision-making. The identification of these novel prognostic markers, rooted in lysosomal gene activity, offers valuable insights for subsequent research and could aid in advancing new risk stratification approaches for MM.

Data availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

Acquah, M. E., Hsing, A. W., McGuire, V. et al. (2019) Presentation and survival of multiple myeloma patients in Ghana: a review of 169 cases. *Ghana Med. J*. **53**, 52-58.

Amaravadi, R. K., Kimmelman, A. C., Debnath, J. (2019) Targeting autophagy in cancer: recent advances and future directions. *Cancer Discov*. **9**, 1167-1181.

Carlson, M., Falcon, S., Pages, H. et al. (2016) hgu133plus2.db: Affymetrix Human Genome U133 Plus 2.0 Array annotation data (chip hgu133plus2). *R package version* **3**.

Cowan, A. J., Green, D. J., Kwok, M. et al. (2022) Diagnosis and management of multiple myeloma: a review. *JAMA* **327**, 464-477.

Davis, S., Meltzer, P. S. (2007) GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* **23**, 1846-1847.

Deng, X., Liu, B., Jiang, Q. et al. (2024) CREBH promotes autophagy to ameliorate NASH by regulating Coro1a. *Biochim. Biophys. Acta Mol. Basis Dis*. **1870**, 166914.

Fonseca, R., Abouzaid, S., Bonafede, M. et al. (2017) Trends in overall survival and costs of multiple myeloma, 2000–2014. *Leukemia* **31**, 1915-1921.

Gray, K. A., Seal, R. L., Tweedie, S. et al. (2016) A review of the new HGNC gene family resource. *Hum. Genomics* **10**, 1-9.

Hanamura, I., Huang, Y., Zhan, F. et al. (2006) Prognostic value of cyclin D2 mRNA expression in newly diagnosed multiple myeloma treated with high-dose chemotherapy and tandem autologous stem cell transplantations. *Leukemia* **20**, 1288-1290.

Heider, M., Nickel, K., Högner, M. et al. (2021) Multiple myeloma: molecular pathogenesis and disease evolution. *Oncol. Res. Treat*. **44**, 672-681.

Heuck, C., Qu, P., Van Rhee, F. et al. (2014) Five gene probes carry most of the discriminatory power of the 70-gene risk model in multiple myeloma. *Leukemia* **28**, 2410-2413.

Huang, Y., Huang, J., Zhang, P. et al. (2021) Integrated analysis of hub gene expression in multiple myeloma. *J. BUON* **26**, 2040-2052.

Huang, Y., Yang, F., Zhang, W. et al. (2023) A novel lysosome-related gene signature coupled with gleason score for prognosis prediction in prostate cancer. *Front. Genet*. **14**, 1135365.

Jiang, Y., Zhou, J., Hou, D. et al. (2019) Prosaposin is a biomarker of mesenchymal glioblastoma and regulates mesenchymal transition through the TGF-β1/Smad signaling pathway. *J. Pathol*. **249**, 26-38.

Liberzon, A., Birger, C., Thorvaldsdóttir, H. et al. (2015) The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. **1**, 417-425.

Mahapatra, K. K., Mishra, S. R., Behera, B. P. et al. (2021) The lysosome as an imperative regulator of autophagy and cell death. *Cell. Mol. Life Sci*. **78**, 7435-7449.

Mahindra, A., Hideshima, T., Anderson, K. C. (2010) Multiple myeloma: biology of the disease. *Blood Rev.* **24**, S5-S11.

Mikkilineni, L., Kochenderfer, J. N. (2021) CAR T cell therapies for patients with multiple myeloma. *Nat. Rev. Clin. Oncol*. **18**,71-84.

Miyahara, Y., Takano, S., Sogawa, K. et al. (2022) Prosaposin, tumor-secreted protein, promotes pancreatic cancer progression by decreasing tumor-infiltrating lymphocytes. *Cancer Sci*. **113**, 2548-2559.

Moreau, P., Kumar, S. K., San Miguel, J. et al. (2021) Treatment of relapsed and refractory multiple myeloma: recommendations from the International Myeloma Working Group. *Lancet Oncol*. **22**, e105-e118.

Ostendorf, T., Zillinger, T., Andryka, K. et al. (2020) Immune sensing of synthetic, bacterial, and protozoan RNA by toll-like receptor 8 requires coordinated processing by RNase T2 and RNase 2. *Immunity* **52**, 591-605.e596.

Padala, S. A., Barsouk, A., Barsouk, A. et al. (2021) Epidemiology, staging, and management of multiple myeloma. *Med. Sci. (Basel)* **9**, 3.

Pan, Y., Meng, Y., Zhai, Z. et al. (2021) Identification of a three-gene-based prognostic model in multiple myeloma using bioinformatics analysis. *PeerJ* **9**, e11320.

Pu, J., Guardia, C. M., Keren-Kaplan, T. et al. (2016) Mechanisms and functions of lysosome positioning. *J. Cell Sci*. **129**, 4329-4339.

Rizvi, A. A., Karaesmen, E., Morgan, M. et al. (2019) gwasurvivr: an R package for genome-wide survival analysis. *Bioinformatics* **35**, 1968-1970.

Robles, J., Pintado-Berninches, L., Boukich, I. et al. (2022) A prognostic six-gene expression risk-score derived from proteomic profiling of the metastatic colorectal cancer secretome. *J. Pathol. Clin. Res*. **8**, 495-508.

Romao, S., Gannage, M., Münz, C. (2013) Checking the garbage bin for problems in the house, or how autophagy assists in antigen presentation to the immune system. *Semin. Cancer Biol*. **23**, 391-396.

Shah, U. A., Mailankody, S. (2020) Emerging immunotherapies in multiple myeloma. *BMJ* **370**, m3176.

Sharma, G., Ojha, R., Noguera-Ortega, E. et al. (2022) PPT1 inhibition enhances the antitumor activity of anti-PD-1 antibody in melanoma. *JCI Insight* **7**, e133225.

Shen, H. M., Mizushima, N. (2014) At the end of the autophagic road: an emerging understanding of lysosomal functions in autophagy. *Trends Biochem. Sci*. **39**, 61-71.

Shi, B., Huang, Q. Q., Birkett, R. et al. (2017) SNAPIN is critical for lysosomal acidification and autophagosome maturation in macrophages. *Autophagy* **13**, 285-301.

Tang, T., Yang, Z. Y., Wang, D. et al. (2020) The role of lysosomes in cancer development and progression. *Cell Biosci*. **10**, 131.

Tayebi, N., Lopez, G., Do, J. et al. (2020) Pro-cathepsin D, prosaposin, and progranulin: lysosomal networks in Parkinsonism. *Trends Mol. Med*. **26**, 913-923.

van Kasteren, S. I., Overkleeft, H. S. (2014) Endo-lysosomal proteases in antigen presentation. *Curr. Opin. Chem. Biol*. **23**, 8-15.

Wang, S., Su, W., Zhong, C. et al. (2020) An eight-circRNA assessment model for predicting biochemical recurrence in prostate cancer. *Front. Cell Dev. Biol*. **8**, 599494.

Wang, T., Qin, Y., Ye, Z. et al. (2023) A new glance at autophagolysosomal-dependent or -independent function of transcriptional factor EB in human cancer. *Acta Pharmacol. Sin*. **44**, 1536-1548.

Wilkerson, M. D., Hayes, D. N. (2010) ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**, 1572-1573.

Yamamoto, K., Venida, A., Yano, J. et al. (2020) Autophagy promotes immune evasion of pancreatic cancer by degrading MHC-I. *Nature* **581**, 100-105.

Yoshihara, K., Shahmoradgoli, M., Martínez, E. et al. (2013) Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun*. **4**, 2612.

Yu, G., Wang, L.-G., Han, Y. et al. (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287.

Zhu, F. X., Wang, X. T., Ye, Z. Z. et al. (2020a) Construction of a prognosis-associated long noncoding RNA-mRNA network for multiple myeloma based on microarray and bioinformatics analysis. *Mol. Med. Rep*. **21**, 999-1010.

Zhu, S. Y., Yao, R. Q., Li, Y. X. et al. (2020b) Lysosomal quality control of cell fate: a novel therapeutic target for human diseases. *Cell Death Dis*. **11**, 817.