

Improving predictive accuracy in multiple myeloma using a plasma cell profile derived from single-cell RNA sequencing

Lanting Liu,^{1,2*} Hao Sun,^{1,2*} Fangshuo Feng,^{1,2} Xiyue Sun,^{1,2} Jingyuan Ma,^{1,2} Rui Lv,^{1,2} Tengteng Yu,^{1,2} Linhai Ye,³ Xiuchun Li,³ Zhen Yu,^{1,2} Xiaoyu Zhang,^{1,2} Huaqing Jing,^{1,2} Yao Yao,^{1,2} Fengxia Ma,^{1,2} Lugui Qiu^{1,2,4} and Mu Hao^{1,2,4}

¹State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Haihe Laboratory of Cell Ecosystem, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Tianjin, China; ²Tianjin Institutes of Health Science, Tianjin, China; ³Sniper Medical Technology Co., Ltd., Suzhou, China and ⁴Gobroad Healthcare Group, Beijing, China

*LL and HS contributed equally as first authors.

Correspondence: M. Hao
haomu@ihcams.ac.cn

Received: February 17, 2025.
Accepted: May 22, 2025.
Early view: May 29, 2025.

<https://doi.org/10.3324/haematol.2025.287586>

©2025 Ferrata Storti Foundation

Published under a CC BY-NC license



Abstract

Multiple myeloma (MM) shows inherent clinical and biological heterogeneity, leading to variable treatment responses and outcomes. The complex molecular landscape of MM makes precise risk stratification through clinical genetic testing difficult. Thus, identifying better biomarkers is essential to enhance existing stratification methods and guide personalized therapy decisions. Here, we systematically analyzed the intratumor heterogeneity of tumor cells from 12 newly diagnosed MM patients with different outcomes at single-cell resolution, especially those with an overall survival of less than 2 years, considered extremely high-risk in the real world. Among the eight heterogeneous tumor cell subclusters in these patients' myeloma cells, a particularly aggressive subset was discovered, characterized by severe chromosomal instability, high-level drug resistance, and high-risk genes. Survival analysis indicated that a high rate of this aggressive cell subset was associated with poor outcomes of the patients. We identified seven genes (*LILRB4*, *CD74*, *TUBA1B*, *CCND2*, *HIST1H4C*, *ITGB7*, and *CRIP1*) with extremely high expression within this subset of aggressive myeloma cells. Multivariate Cox analysis showed that the seven-gene signature score was the worst factor for patients' outcome independently of aberrant cytogenetics and International Staging System stage. We then established an integrated risk stratification model combined with the seven-gene signature score. This model significantly improved the risk discrimination capabilities, especially in distinguishing the ultra-high-risk myeloma patients with the worst outcome in our cohort, and was validated in five independent datasets of MM patients. We further devised a simple digital polymerase chain reaction method for feasible quantification of the seven-gene signature, which still significantly differentiated the survival of MM patients and has considerable value for clinical application. Overall, this integrated risk-scoring model derived from single-cell RNA-sequencing data was significantly associated with a more advanced stage of myeloma, facilitating guided risk-adapted treatment strategies for such ultra-high-risk patients.

Introduction

In the past decades, advancements in therapies and the widespread application of novel drugs and autologous stem cell transplantation have led to improved overall survival rates for patients with multiple myeloma (MM).^{1,2} However, some high-risk patients still struggle to benefit from current treatment strategies.^{3,4} Patients with ultra-high-risk MM are currently defined as those who die within 24-36 months of diagnosis, while high-risk

MM patients are those who die within 36-60 months of diagnosis. Risk stratification and risk-adapted therapy are critical for identifying groups of patients in urgent need of novel therapies. Although there are several validated risk stratification systems in routine clinical practice, a unified approach for defining risk of newly diagnosed MM (NDMM) patients remains elusive.⁵

The Revised International Staging System (R-ISS) integrates interphase fluorescence *in situ* hybridization abnormalities, lactate dehydrogenase (LDH) level, and

features of the International Staging System (ISS) to serve as the predominant risk stratification tool for predicting overall survival and progression-free survival in patients with NDMM.⁶ However, in the R-ISS system, the intermediate-risk group accounts for 60% to 70% of patients, highlighting the need for more refined stratification of this cohort. To address this need, the second revision of the International Staging System (R2-ISS) has included factors such as *TP53* mutations and 1q21 amplification.⁷ Compared to R-ISS, R2-ISS provides better stratification for intermediate-risk MM patients and helps us to assess these patients' prognosis more accurately.

In recent years, gene expression profiling has emerged as a risk stratification tool for MM based on the genomic composition of myeloma cells, offering additional prognostic subgroups. UAMS-70 and SKY92 are both prognostic scoring systems based on gene expression profiles, enabling a more comprehensive and precise evaluation of patients' prognosis, ultimately to guide personalized treatment approaches.^{8,9} However, the increased complexity and demands on resources of the gene expression profiling scoring systems may hinder their widespread adoption compared to the R-ISS.

In the present study, ISS staging, del(17p), 1q+, t(4;14), and the seven-gene score were integrated into a comprehensive weighted myeloma prognostic score system (MR-ISS), which could categorize patients into four risk groups. Compared to ISS, R-ISS, and R2-ISS, the MR-ISS model demonstrated better differentiation and identification of ultra-high-risk patients with an overall survival of less than 24 months.

Methods

Patients' characteristics and single-cell sequencing

Bone marrow mononuclear cells were isolated by Ficoll separation from seven healthy donors and 12 NDMM patients. Written informed consent was obtained from all participants prior to sample collection. The clinical characteristics of the patients are provided in the *Online Supplementary Materials*. This study was conducted in accordance with the principles of the Declaration of Helsinki and approved by the Ethics Committee of the Institute of Hematology and Blood Diseases Hospital, Chinese Academy of Medical Sciences (China) (protocol code KT2020010-EC-2). Single-cell RNA libraries were constructed using Chromium Single Cell 3' Reagent Kits (v2 Chemistry) and sequenced on an MGISEQ-2000 sequencer by the Beijing Genomics Institute (BGI, China).¹⁰⁻¹³

Statistical analysis

Data analyses were performed with R language and GraphPad Prism 8.0 software. Statistical significance was set at $P < 0.05$.

Results

Investigation of the heterogeneous component of tumor cells between patients who died within or beyond 24 months

In order to clarify the intrinsic molecular characteristics of MM cells in those patients with an overall survival of less than 2 years, considered ultra-high-risk in the real world, we conducted single-cell transcriptome profiling on bone marrow plasma cells (MM cells) from 12 NDMM patients with varied clinical characteristics, different cytogenetic abnormalities, and diverse outcomes. We compared the profiles of the MM cells with those of cells obtained from seven healthy donors. According to the follow-up data, four individuals (MM8, MM12, MM15 and MM24) among the 12 patients who exhibited rapid disease progression and early mortality within 24 months were classified as clinically ultra-high-risk and designated as the EM24 group, the others (MM1, MM2, MM3, MM4, MM5, MM16, MM25, and MM27) were designated as the non-EM24 (nEM24) group. Among EM24 patients, routine fluorescence *in situ* hybridization assay demonstrated that four (4/4) patients harbored 1q(+), one (1/4) patient had 17p(-), one (1/4) patient had t(4;14) and two (2/4) patients had elevated LDH. The M paraprotein in these cases was IgD, IgA, IgG and light chain, respectively. In the nEM24 group, four (4/8) patients had 1q(+), three (3/8) patients had t(4;14), but none of them had elevated LDH and 17p(-). Twelve patients were risk-stratified according to ISS and R-ISS criteria, and received standard bortezomib, lenalidomide, and dexamethasone (VRD) combination therapy. Detailed clinical and pathological information on the MM patients are summarized in *Online Supplementary Table S1*.

First, we conducted principal component analysis to reveal the differences in cell clustering between healthy donors, and EM24 and nEM24 patients at single-cell resolution (Figure 1A). As expected, the plasma cells of MM patients exhibited a marked increase in the component of myeloma/plasma cells (Figure 1B). In order to efficiently identify intrinsic heterogeneity in MM cells between the patients in the EM24 and nEM24 groups, the analysis integrated MM/plasma cell compartments from all samples into a combined dataset. Eight distinct subpopulations (sub-C0-C7) were identified with high expression of plasma cell marker genes, *CD38*, *SDC1*, and *TNFRSF17* (Figure 1C, D). The tumor cell compositions of the EM24 MM patients were significantly different from those of the nEM24 patients. Of note, among the sub-C0-C7 subclusters, sub-C4 was a unique, significantly increased subpopulation in EM24 MM patients compared to nEM24 patients (Figure 1E). Odds ratio analysis further confirmed that sub-C4 exhibited a particularly strong distribution preference among EM24 patients (Figure 1F). These findings strongly suggest that the presence of sub-C4 MM cells plays a critical role in the aggressive progression of MM in EM24 patients.

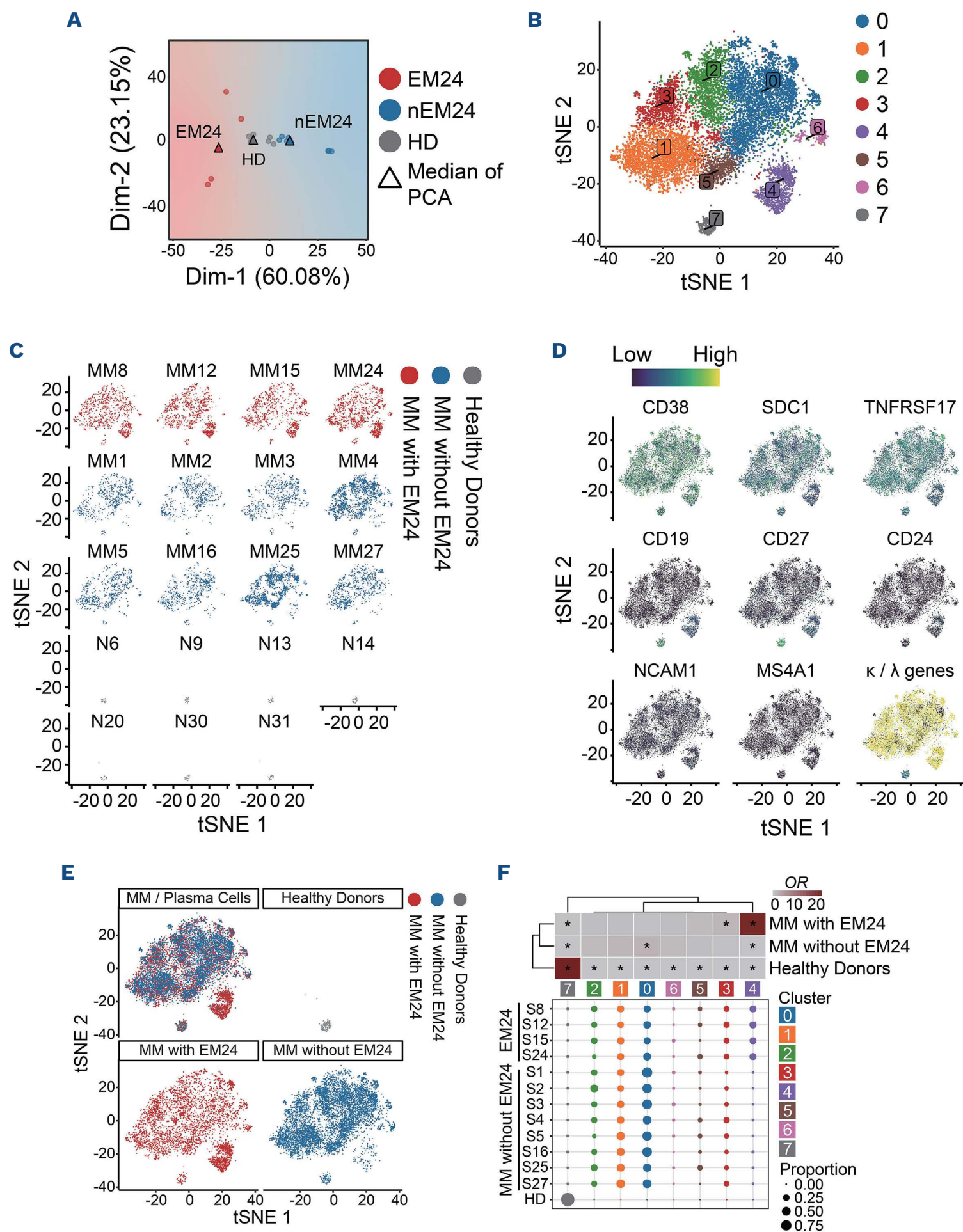


Figure 1. Single-cell transcriptomic profiling of the bone marrow ecosystem in patients with multiple myeloma. (A) Two-dimensional plots showing the distribution of each sample. The color of the points indicates the sample group. Triangles represent the median of the principal component analysis (PCA). PCA was used to assess differences in cell clustering at single-cell resolution across healthy donors, multiple myeloma (MM) patients who died early (EM24), and MM patients who died later (nEM24). (B) T-distributed stochastic neighbor embedding plots showing the distribution of annotated plasma cells. Cell annotations are labeled by colors. (C) Uniform manifold approximation and projection (UMAP) plot showing ten plasma cell clusters from all samples. (D) UMAP plots showing the expression of marker genes and immunoglobulin light chain κ/λ ratio in plasma cells. (E) UMAP plots showing the plasma cell origins of the three groups of samples. (F) Heatmap illustrating the odd ratios of each cluster in each sample group based on the Fisher exact test (top). Asterisks represent $P < 0.05$. The bar plot shows the proportion of each cluster in each MM sample and healthy donors (bottom). EM24: patients who died early; nEM24: patients who died later than those in the EM24 group; HD: healthy donors; tSNE: t-distributed stochastic neighbor embedding; OR: odds ratio.

Heterogeneous molecular characteristics in different myeloma cell clusters

Myeloma cells exhibit significant intrinsic heterogeneity. Elucidating the molecular heterogeneity of MM cells could be instrumental in identifying novel biomarkers for the discrimination of ultra-high-risk patients and facilitating therapeutic improvements for the treatment of such patients. Gene expression profiling revealed uniquely expressed genes and differentially expressed genes (DEG) within the eight tumor cell subpopulations (C0 to C7). The top five DEG in each plasma cell subcluster are displayed in a heatmap (Figure 2A). Moreover, severe chromosomal instability was observed in sub-C4 MM cells, with various gains and losses of chromosome segments, including del(1p), del(2q), del(10q), del(13), del(14), del(17p), amp(1q), amp(12p), and amp(19) (Figure 2B). These findings demonstrate that MM cells in sub-C4 constitute a subpopulation of MM cells with an aggressive phenotype including activated proliferation. Copy number variation scores were utilized to evaluate the copy number variation in each cluster of MM cells. The data show that the copy number variation score was highest in the fourth subcluster (Figure 2C), indicating that there is severe chromosomal instability of the cells in sub-C4 among all subclusters.

Previous studies reported a 70 high-risk gene set (UAMS-70) and a 56 drug-resistance gene set associated with inferior outcomes in MM. We, therefore, next examined the expression of those malignant genes within the heterogeneous subpopulations of MM cells. Strikingly, our results clearly showed that, among all MM cells, sub-C4 MM cells expressed the highest levels of high-risk and drug resistance genes (Online Supplementary Figure S1A, B). Of note, 84.3% (59/70) of high-risk genes and 55.3% (31/56) of drug-resistance genes showed increased expression in sub-C4 MM cells. We found that the PRC2-associated factor, PHF19 (*PLC3*), which promotes the proliferation and drug-resistance of MM cells as we reported previously,¹⁴ was overexpressed in sub-C4 cells. These findings further support that sub-C4 encompasses aggressive MM malignant features.

Identification of the specific gene expression characteristics and clinical impact of sub-C4 myeloma cells

We next compared the gene expression of sub-C4 cells to that of the other subclusters. We found 1,320 DEG ($\log_{FC} > 0.25$, $P < 0.01$), 1,263 upregulated and 57 downregulated, after excluding immunoglobulin-associated genes (Figure 3A). The top ten upregulated genes in sub-C4 compared with other subclusters are shown in Online Supplementary Table S2. This profile included *LILRB4*, *STMN1*, *GAPDH*, *CRIP1*, *TUBA1B*, *TUBB*, *HIST1H4C*, *CD74*, *ITGB7* and *CCND2*. Our results indicated that *LILRB4*, encoding an immunosuppressive receptor, was significantly upregulated in sub-C4 and expressed at very high levels ($\text{avg_log}_{FC} > 1.5$, $P < 0.001$). To further elucidate the gene expression regulatory network

and the biological characteristics associated with DEG in sub-C4 tumor cells, we conducted a comprehensive network analysis to identify key transcription factors and their downstream targets within these DEG. Additionally, we employed L1 regularization analysis to investigate specific genomic alterations in sub-C4, as identified through transcription factor gene networks and gene co-expression networks. The results revealed that MYC and E2F1, which were highly activated within the sub-C4 population, acted as key transcription factors promoting tumor growth by transcriptionally regulating downstream gene expression (Figure 3B). Notably, hub genes including *LILRB4*, *CD74*, *XBP1*, *MKI67*, *AURKB*, and *USP1* were found to be closely related to the sub-C4 phenotype (Figure 3C). Subsequently, we performed gene ontology enrichment analysis on the sub-C4 cluster and found that its main functions are concentrated in 'Lymphocyte-mediated immunity', 'Adaptive immune response based on somatic recombination of immune receptors', 'Immunoglobulin-mediated immune response', 'Antigen receptor-mediated signaling pathway' and 'B cell-mediated immunity' (Figure 3D). Subsequently, we conducted an analysis to estimate the composition of the eight myeloma/plasma cell subclusters (C0-C7) in the GSE2658 dataset (N=414). Survival analysis revealed a significant association between a high proportion of sub-C4 plasma cells and poor prognosis among patients (Figure 3E). These findings indicate that sub-C4 plasma cell abundance is a significant prognostic marker in MM patients, providing strong evidence for predicting unfavorable outcomes in NDMM patients. Among the top ten upregulated genes, univariate Cox regression analysis showed that *LILRB4*, *ITGB7*, *CRIP1*, *TUBA1B*, *CCND2*, *HIST1H4C* and *CD74* were significantly correlated with inferior outcomes of MM patients, based on the GSE2658 dataset of NDMM patients (Figure 3F).

The seven-gene signature facilitates discrimination of multiple myeloma patients' outcome

Our study further investigated the power of stratification efficacy of a seven-signature model in discriminating ultra-high-risk MM patients. Based on Cox regression analysis, we developed a prognostic risk score formula using the expression levels of seven genes. Each patient was assigned a risk score, with the optimal cut-off determined by maximally selected rank statistics. The high-risk MM patients with R-ISS stage III from the GSE136324 dataset (training cohort) could be further classified into two groups: a high-risk group (blue) and an ultra-high-risk group (red). Kaplan-Meier analysis showed that patients in the ultra-high-risk group had significantly shorter overall survival than those in the high-risk group (Figure 4A). To confirm the reliability of the seven-gene signature, the bulk RNA-sequencing data of CD138⁺ cells from bone marrow of MM patients in our center with high risk identified by mSMART 3.0 were utilized as a validation cohort. The risk score for each patient in this validation co-

hort was calculated using the formula mentioned above. Our data showed that the seven-gene signature could further segregate ultra-high-risk patients utilizing the best cut-off point of the risk score. Consistently, patients in the ultra-high-risk group had shorter overall survival than those in the high-risk group (Figure 4B). Therefore, these data indicate that the seven-gene signature established by our study is

an extremely powerful model in further risk stratification of ultra-high-risk patients as identified by either R-ISS stage III or mSMART 3.0.

We next compared the efficiency of our model with that of the previously established SKY-92 and USMA70 models using the same GSE136337 dataset, which includes 66 high-risk MM patients with R-ISS stage III. Notably, the results

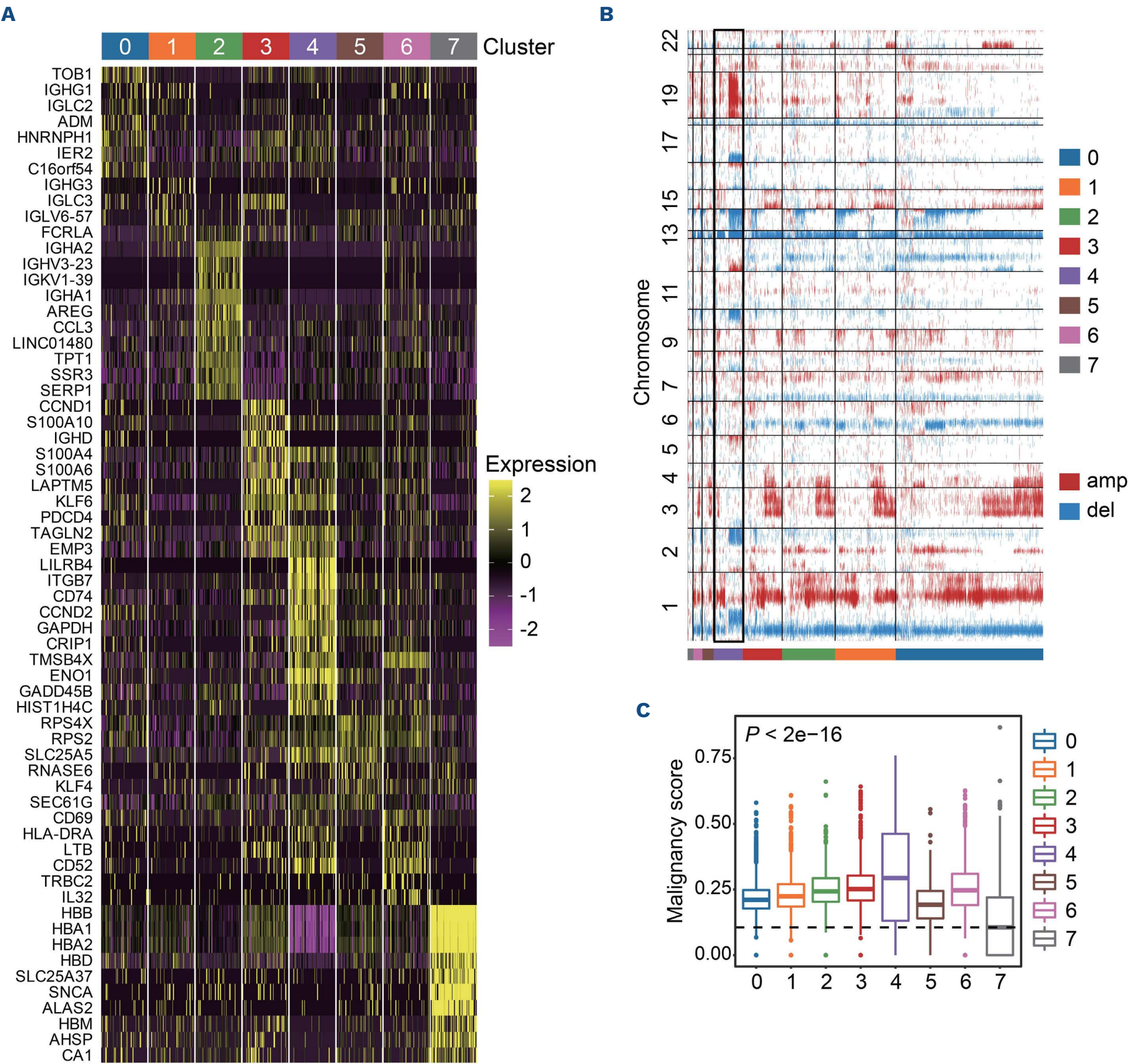


Figure 2. Comparative gene expression profiling of sub-cluster populations of myeloma cells. (A) Heatmap showing the top ten differentially expressed genes among diverse sub-clusters of myeloma cells. The sub-cluster number is labeled on the top of graph. (B) The landscape of inferred large-scale single cell copy number variations for all sub-clusters. The annotation track on the right corresponds to the sub-clusters. Chromosome numbers are labeled at the left. Blue indicates deletion of chromosomes; red indicates amplification of chromosomes. (C) The malignancy score across multiple myeloma cell clusters. The malignancy score is calculated based on copy number variation scores of each cell. Amp: amplification; del: deletion.

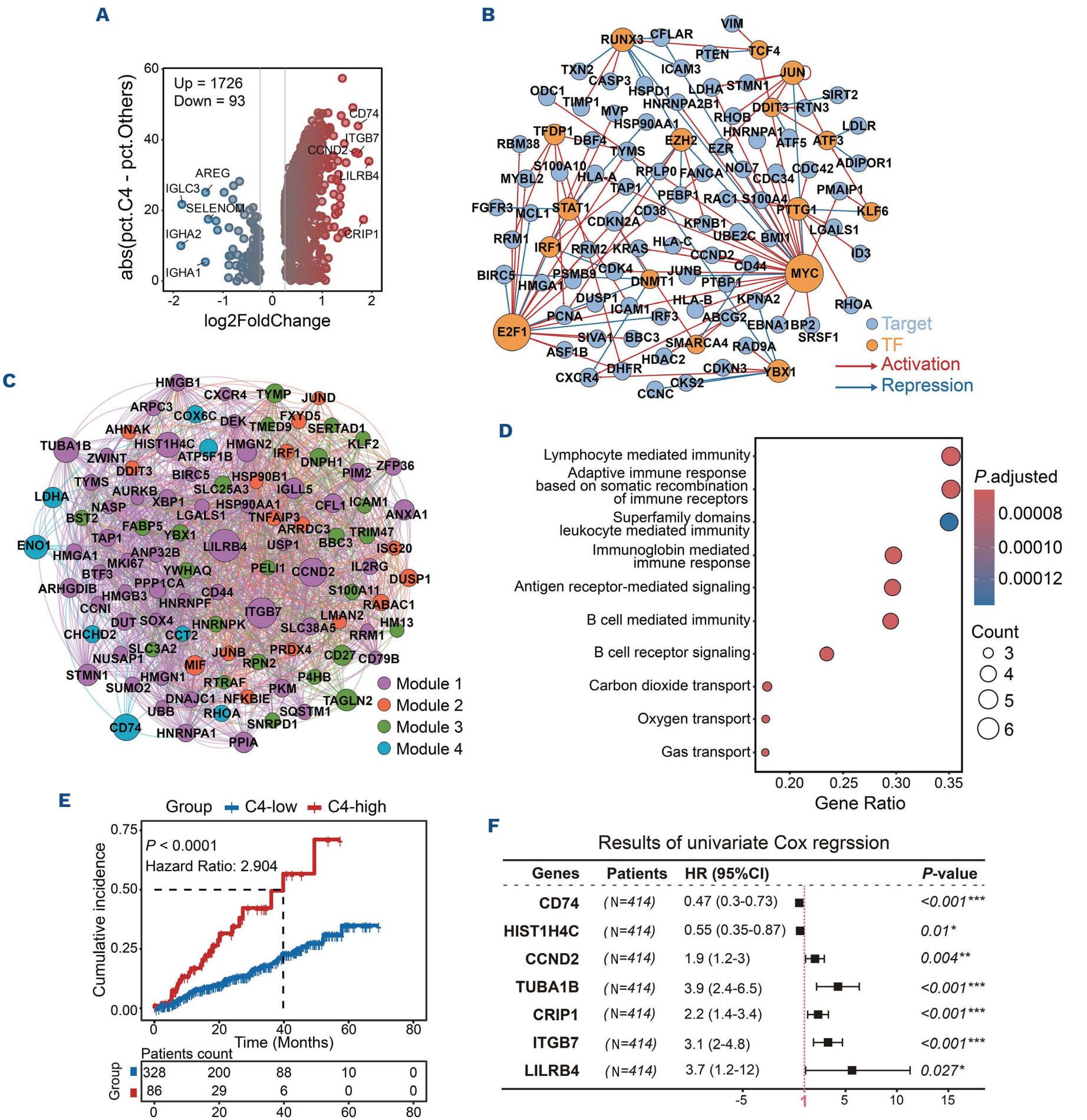


Figure 3. Identification of genes associated with a high risk in sub-C4 plasma cells. (A) Volcano plot showing differentially expressed genes (DEG) between sub-C4 and the rest of the sub-clusters. Upregulated genes are indicated as red dots and downregulated genes are indicated as blue dots. The names of the most significant DEG are marked in the plot. (B) Regulatory network of major transcription factors and target genes upregulated in sub-C4 compared with the other multiple myeloma (MM) cells. The point size indicates the edge numbers of transcription factors. Red lines indicate active signal pathways, blue lines indicate suppressive pathways. (C) Gene co-expression network displaying the top gene modules specific to subcluster 4. The point size indicates the gene fold change in sub-C4 versus the other MM cells. The lines connect significantly correlated genes. Genes of the same color represent interrelatedness. (D) Gene ontology term enrichment analysis for sub-C4 according to the DEG between sub-C4 and the rest of the subclusters. (E) Kaplan-Meier curve showing the overall survival of 414 MM patients in the GSE2658 dataset according to whether they had a high or low proportion of subcluster 4 cells. A log-rank test was applied for the comparison between groups. (F) Forest plot showing the univariate Cox regression analysis of seven genes in patients from the GSE2658 dataset. TF: transcription factor; HR: hazard ratio; 95% CI: 95% confidence interval. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

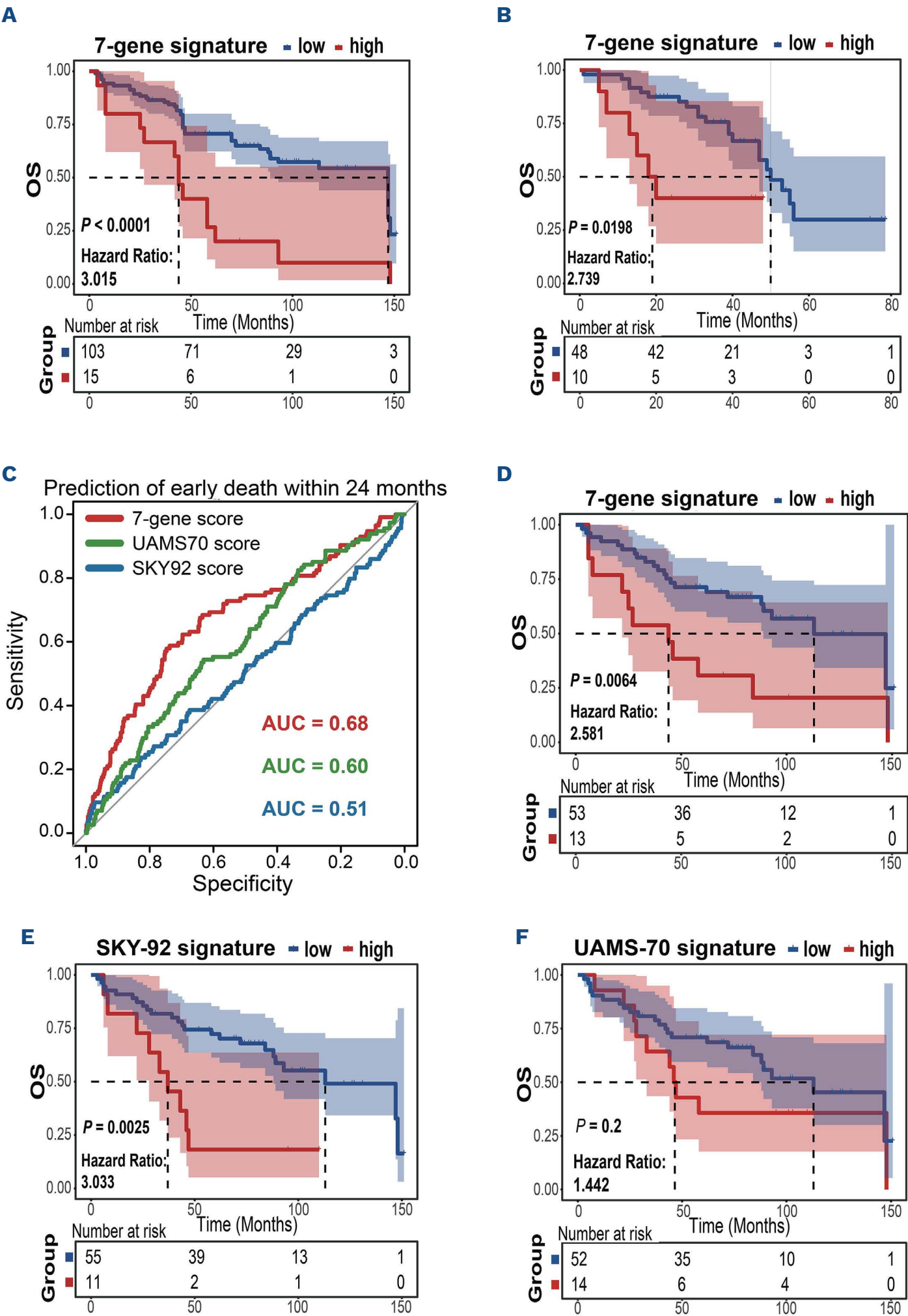


Figure 4. A seven-gene signature facilitates identification of ultra-high-risk myeloma patients. (A, B) The Kaplan-Meier analysis of the seven-gene signature in the GSE136324 training cohort (left) and in the validation cohort (in-house bulk RNA-sequencing dataset, right). (C) Receiver operating characteristic plots showing the sensitivity and specificity of the seven-gene score, UAMS70 score, and SKY92 score in predicting multiple myeloma patients who will die early. (D-F) Kaplan-Meier analysis with the seven-gene signature (D) and the SKY-92 signature (E) in the GSE136337 dataset and the UAMS-70 signature (F) in high-risk multiple myeloma patients with Revised International Staging System stage III disease in the GSE136337 dataset. OS: overall survival; AUC: area under the curve.

showed that the prognostic efficacy of the seven-gene signature was equal to that of SKY-92 and superior to that of UAMS-70 (Figure 4C-F). Moreover, our model includes seven genes, so it would be easier to use in clinical practice.

Exhausted T cells with metabolic dysfunction were identified in multiple myeloma patients who died within 24 months

According to our finding indicated in Figure 3D, sub-C4 function is enriched in immune-related signal pathways. To better identify the crosstalk between sub-C4 MM cells and T-cell subsets, we generated a deep transcriptional atlas of T cells in the tumor microenvironment of multiple cancers from 20 human samples. We defined five functional CD4⁺ and seven CD8⁺ T-cell clusters based on canonical patterns of marker expression across conditions (*Online Supplementary Figure S2A-E*). We made a projection of our single-cell sequencing data onto the T-cell atlas and identified the various T-cell clusters in our cohort (Figure 5A). To understand the extent to which the immune microenvironment contributed to the rapid MM progression, we compared the composition of T-cell clusters among EM24 MM patients, nEM24 MM patients and healthy donors. EM24 MM patients demonstrated increased numbers of exhausted-like and immune-suppressive T cells, including CD8-TPEX, CD8-TEX, CD4-Exh, CD4-Th17, and Treg (Figure 5B, C).

Combining with the characteristic phenotypic features and the calculated naïve, cytotoxic, and exhausted scores, the canonical trajectories of naïve T cells were identified to follow either effector-memory fate or exhausted fate (*Online Supplementary Figure S2F-J*). Effector-memory T cells (cell fate 2) were significantly active in T-cell-mediated cytotoxicity, T-cell activation, WNT signaling, MAPK6/MAPK4 signaling, IL-1-induced activation of NF- κ B, the CD40 signaling pathway, and mTOR signaling. By contrast, exhausted T cells (cell fate 1) were active in IFN- γ signaling, alternative mRNA splicing, and the PD-1 pathway (Figure 5D). Consistent with our previous study, a significant distinction existed in metabolic state between effector-memory T cells and exhausted T cells. Obviously, exhausted T cells exhibited lower levels in glycolysis, the tricarboxylic acid cycle, glutamine metabolism and lipid metabolism compared with effector-memory T cells (Figure 5E, *Online Supplementary Figure S2K, L*). Additionally, analysis of transcription factors revealed significant upregulation of JUN, FOS, and NR4A1 in effector-memory T cells and upregulation of EMOES, TOX, and HOPX in exhausted T cells. As expected, exhausted T cells were characterized by higher expression of immune checkpoints (*Online Supplementary Figure S2K*).

The evident exhausted T cells in EM24 MM patients raised the question of whether the memory T cells in EM24 patients had already been in an aberrant state. To answer this question we performed DEG analysis of memory T cells between EM24 patients and nEM24 patients (*Online Supplementary*

Figure S2M, N). We found that the genes upregulated in memory T cells from EM24 patients were enriched in apoptotic signaling pathway, histone modification, IL-4 and IL-13 signaling, oxidative phosphorylation, and oxidative stress-induced senescence. Formation of ATP, T-cell differentiation, and co-stimulation by the CD28 family were significantly downregulated pathways in memory T cells from EM24 patients (*Online Supplementary Figure S2O*). Collectively, the enrichment of exhausted T cells was characterized by aberrant metabolism and high expression of immune suppressive checkpoints. On the other hand, the memory T cells committed to the generation of effector cells showed defects in differentiation capability and increased apoptotic signaling. These results indicate that patients in the EM24 group exhibit MM cells with more aggressive characteristics, and their immune microenvironment is significantly more dysfunctional compared to that of nEM24 patients. The crosstalk between sub-C4 MM cells and T-cell subsets plays a crucial role in establishing an immunosuppressive tumor microenvironment in EM24 patients.

Investigations on the impact of integrating the seven-gene signature and clinical characteristics in the identification of ultra-high-risk multiple myeloma patients

Given the biological characteristics of sub-C4 cells and their crosstalk with T cells in the immunosuppressive tumor microenvironment, an attempt was made to improve the risk stratification of MM patients by integrating the seven-gene signature of sub-C4 MM cells and clinical features. We incorporated existing MM high-risk factors into univariate and multivariate Cox regression analysis of the MMRF-CoMMpass dataset as the training cohort, and GSE136337 and in-house datasets as the validation cohorts. Age, gender, ISS stage, del(17p), 1q+, t(4;14), t(14;16), LDH level, and the seven-gene signature were included (*Online Supplementary Table S3*). Univariate and multivariate Cox analyses were conducted (Figure 6A, B), which showed that ISS stage, del(17p), 1q+, t(4;14) and the seven-gene signature, but not t(14;16) or LDH level, were independent variables for predicting the overall survival of MM patients. Based on this additive score system, we developed a modified R-ISS (MR-ISS) that accurately stratifies patients into four groups: low risk (MR-ISS I, 0 points), intermediate risk (MR-ISS II, 1-1.5 points), high risk (MR-ISS III, 2-4 points), and ultra-high risk (MR-ISS IV, 4.5-6 points) (Figure 6C). We compared the performance of different staging systems for risk stratification within the same cohort of patients. In particular, both the MR-ISS and R2-ISS showed improved discriminatory capability among intermediate-risk patients compared to R-ISS (Figure 6D). However, the ISS, R-ISS, and R2-ISS could not discriminate among patients whose median overall survival was less than 2 years (Figure 6E, F, *Online Supplementary Figure S3A*).

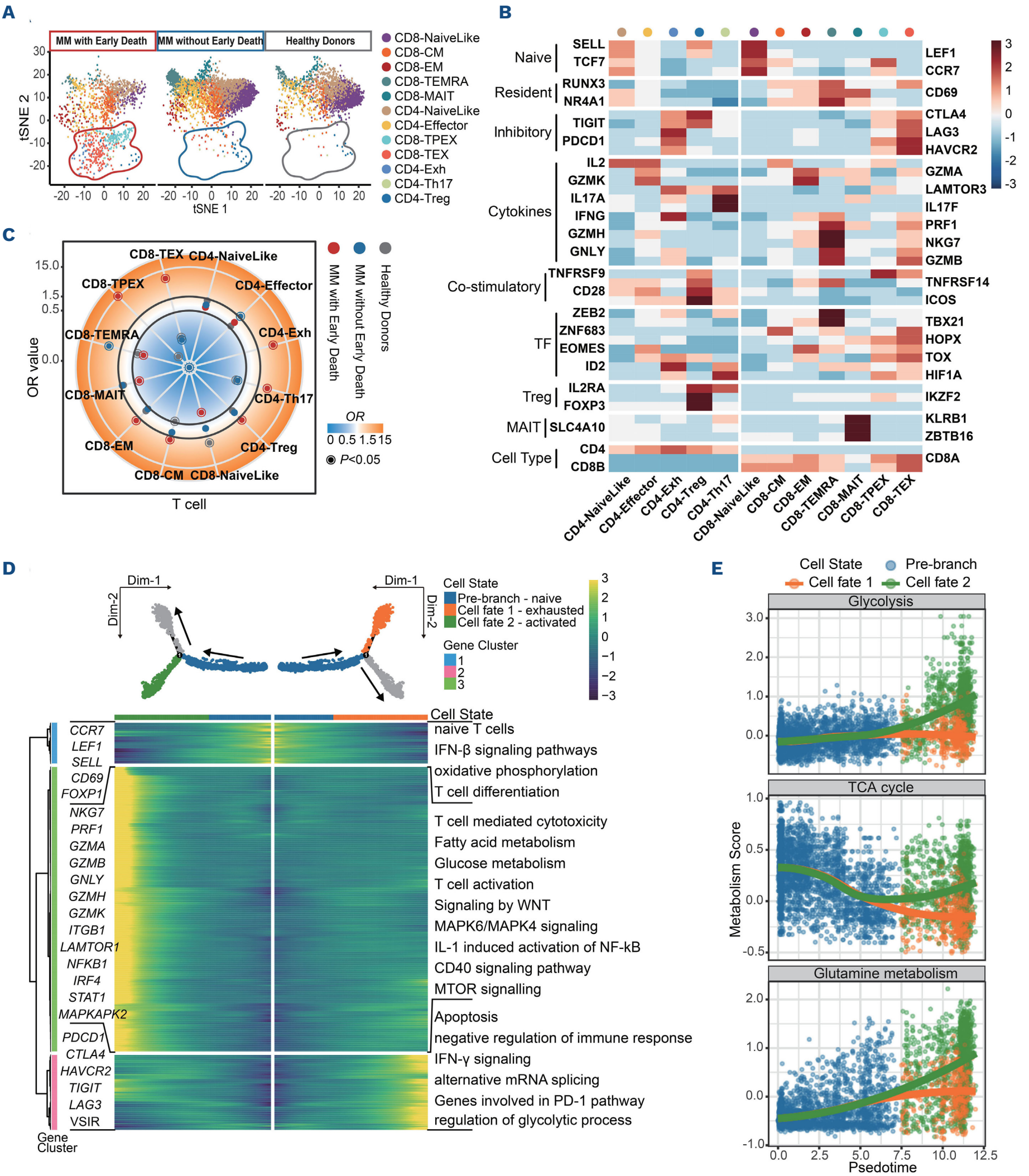


Figure 5. Metabolic dysfunction in exhausted T cells was observed in ultra-high-risk myeloma patients. (A) Uniform manifold approximation and projection plots showing the distribution of CD8⁺ T and CD4⁺ T cells across three groups of samples. Left: samples from patients who died early; middle: non-early death patients; right: healthy donors. (B) Dot plot illustrating the odds ratio (OR) of each T-cell subset proportion across the three groups of samples. Dot colors represent sample groups. OR values

Continued on following page.

are based on the Fisher exact test. A high OR indicates that the cell is more likely to distribute in the group, while a low OR indicates that the cell is less likely to distribute in the group. (C) Heatmap depicting the expression of representative gene sets across T-cell subtypes, including naïve, resident, inhibitory, cytokine-related, co-stimulatory, transcription factors, regulatory T cells, mucosal-associated invariant T cells, and other cell types. (D) Pseudo-time trajectory defining the cell states of CD8⁺ T cells (upper panel). Heatmap showing dynamic changes in gene expression along the pseudo-time trajectory (grouped into three clusters) and their enriched biological pathways (lower panel). (E) Fitted line with dot plot showing metabolism scores along pseudo-time. Colors represent different states of T cells. tSNE: t-distributed stochastic neighbor embedding; MM: multiple myeloma; CM: central memory cells; EM: effector memory cells; TEMRA: terminally differentiated effector memory cells; MAIT: mucosal-associated invariant T cells; TPEX: progenitor exhausted T cells; TEX: terminally exhausted T cells; Exh: exhausted cells; T reg: regulatory T cells; TF: transcription factor; TCA: tricarboxylic acid.

The Modified Revised International Staging System effectively improves the ability to distinguish ultra-high-risk patients with significantly poorer outcome

To further validate the MR-ISS model in risk stratification, publicly available datasets of MM cohorts were used. In the MMRF-CoMMpass cohort, we found that the MR-ISS model provided excellent discrimination of MM patients with diverse outcomes, especially the ultra-high-risk patients with overall survival of less than 24 months ($P < 0.001$) (Figure 7A). In the independent external validation cohort, the median overall survival was 165, 146, 93 and 40 months in the MR-ISS stage I, II, III, and IV groups, respectively ($P < 0.001$) (Figure 7B). In addition, we utilized an in-house cohort for internal validation, which confirmed that the median overall survival was not reached, 71.0, 41.4, and 24.9 months for MR-ISS stage I, II, III, and IV, respectively ($P < 0.001$) (Figure 7C). Notably, the median overall survival of MR-ISS stage IV NDMM patients was less than 25 months in both the MMRF-CoMMpass and in-house datasets, indicating that the MR-ISS established in this study improved the ability to distinguish the ultra-high-risk patients with significantly poorer outcome.

The droplet digital polymerase chain reaction (ddPCR) technique is a highly sensitive and precise method for quantifying nucleic acids.¹⁵ To further establish a clinically feasible approach to biomarker detection for guiding treatment selection and prognosis stratification in MM patients, we devised a ddPCR method for quantifying seven genes in 139 newly diagnosed patients and combined it with the risk factors of MR-ISS for scoring. The findings indicated that the seven-gene signature determined by the ddPCR quantification method still significantly differentiated the survival of MM patients and has considerable value for clinical application (stage IV vs. I: $P < 0.0001$) (Figure 7D) (stage IV vs. I: $P < 0.05$) (Figure 7E). By integrating these two approaches, clinicians can make more informed decisions regarding personalized treatment strategies for MM patients.

Discussion

MM is a plasma cell malignancy characterized by heterogeneity of tumor cells, which influences the outcome of patients.¹⁶ There have been many models for risk stratification of MM, but it is challenging to identify the ultra-high-risk

patients with an overall survival of less than 24 months.^{6,7,17} Much research has focused on whether gene expression profiles can enhance the accuracy of risk stratification, including the UAMS70 and SKY92 models.^{18–21} However, they are based on the DEG of MM cells, which cannot reflect the heterogeneity between different MM cells. In this study, single-cell RNA-sequencing was conducted to investigate the heterogeneous characteristics among MM cells, and to relate differences in tumor cells between patients with distinct outcomes. Based on a comprehensive molecular characterization of tumor cells and immune cells in the tumor microenvironment, this study revealed the entire spectrum of the bone marrow ecosystem in MM, thereby providing the basis for a new risk stratification model for the identification of ultra-high-risk patients. The MR-ISS risk stratification system that integrates clinical features, genomic aberrations, and changes in gene expression of MM cells. Therefore, this system provides a more comprehensive reflection of the tumor biology and clinical characteristics, thus achieving a more refined clinical risk stratification. The MR-ISS risk stratification system also provides the basis that individual treatment is required for ultra-high-risk patients. Beside the aggressive biological features of tumor cells, our study further confirms that the crosstalk between tumor cells and tumor microenvironment plays a pivotal role in tumor progression and the outcome of patients. Enrichment of exhausted T cells characterized by aberrant metabolism and high expression of immune suppressive checkpoints was found in the tumor microenvironment of patients in the EM24 group. Moreover, the memory T cells that were committed to the generation of effector cells showed defects in differentiation capability and increased apoptotic signaling. Therefore, the ideal treatment strategy for ultra-high-risk patients should target tumor cells and rescue the immunosuppressive tumor microenvironment. In addition, our findings support the concept that T-cell dysfunction not only involves the stage of effector T cells, but memory T-cell dysfunction should also be rescued. Seven genes, namely *LILRB4*, *CD74*, *TUBA1B*, *CCND2*, *HIST1H4C*, *ITGB7* and *CRIP1*, were found to be highly expressed in the sub-C4 plasma cell population. *LILRB4* (immune evasion/NF- κ B activation),^{22,23} *CD74* (immunosuppression/drug resistance),²⁴ *TUBA1B* (proliferation/microtubule stabilization),²⁵ *CCND2* (cell cycle regulation),²⁶ *HIST1H4C* (proliferation/epigenetics),^{27,28} *ITGB7* (adhesion/metastasis),^{29,30} and *CRIP1*

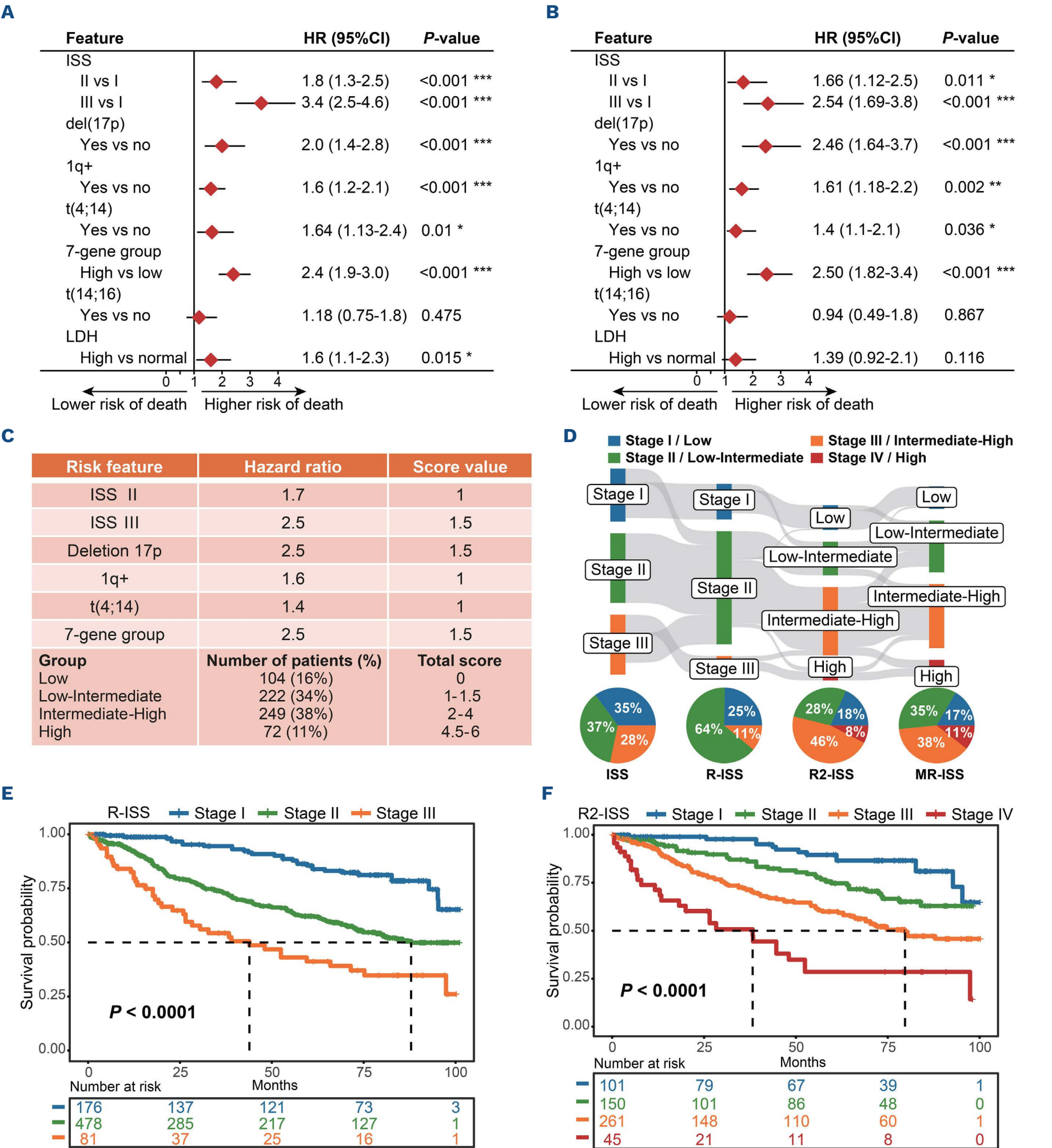


Figure 6. Combining gene expression profiling-based biomarkers with the Revised International Scoring System score. (A) Forest plot showing the univariate Cox regression result of International Scoring System (ISS) stage, del(17p), 1q+, t(4;14), seven-gene group, t(14;16), and lactate dehydrogenase (LDH) level in multiple myeloma (MM) patients in the MMRF-CoMMpass dataset. (B) Forest plot showing the multivariate Cox regression result of ISS stage, del(17p), 1q+, t(4;14), seven-gene group, t(14;16), and LDH level in MM patients in the MMRF-CoMMpass dataset. (C) Modified Revised ISS (MR-ISS) score definition based on the hazard ratio from multivariate Cox regression in MM patients included in the training set. (D) Sankey diagram and pie plots showing the relationship and proportion of MM patients' stages by ISS, Revised ISS (R-ISS), second revision of the ISS (R2-ISS), and the MR-ISS. (E, F) Kaplan-Meier curve showing the overall survival of MM patients in the MMRF-CoMMpass dataset classified by the R-ISS (E) and R2-ISS (F) systems. HR: hazard ratio; 95% CI: 95% confidence interval. **P*<0.05, ***P*<0.01, ****P*<0.001.

(proteostasis/autophagy)³¹ form a critical pathogenic network in MM. These genes collectively drive tumor proliferation, drug resistance, and immune microenvironment dysregulation. *LILRB4* is the top uniquely expressed gene in the sub-C4

population. *LILRB4* is an immunosuppressive receptor that plays a role in regulating the immune response and is involved in the pathogenesis of acute myeloid leukemia. Recently, we and other groups independently reported a high level of *LILRB4* on MM cells and that the receptor

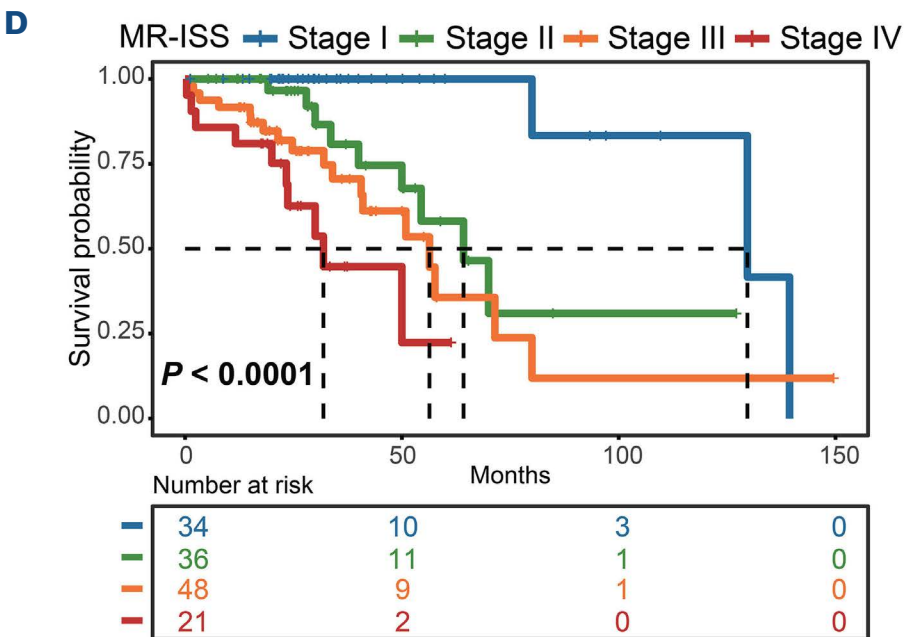
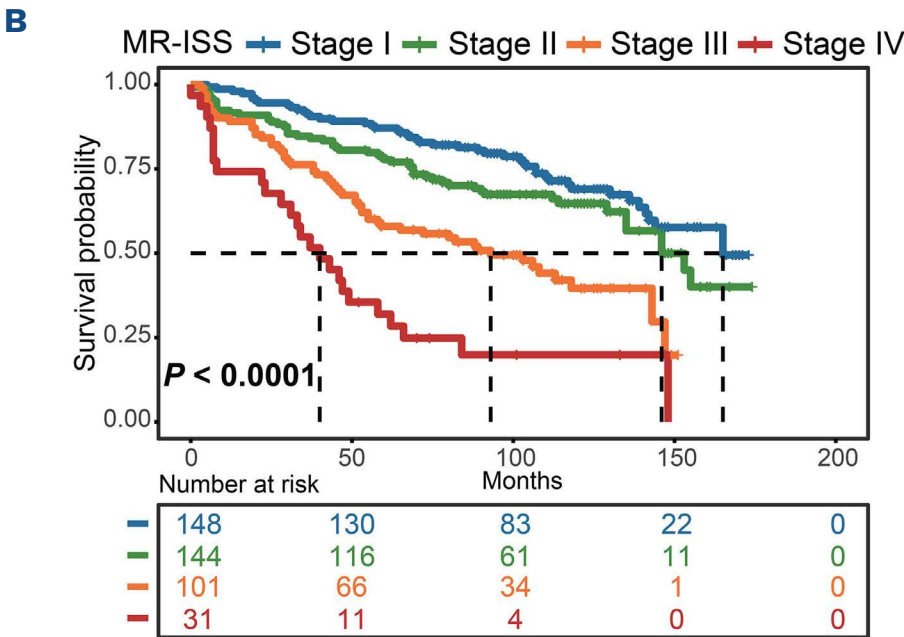
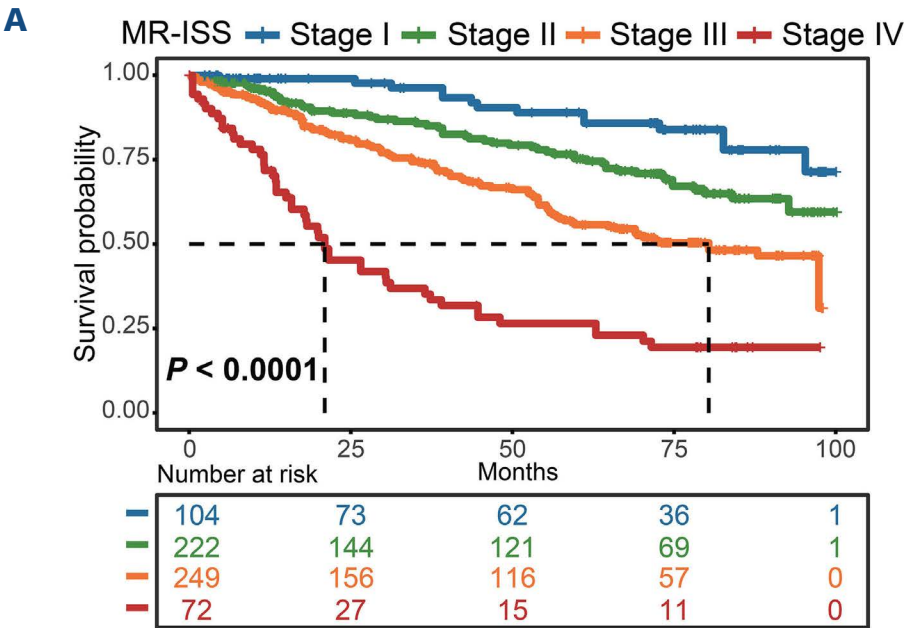
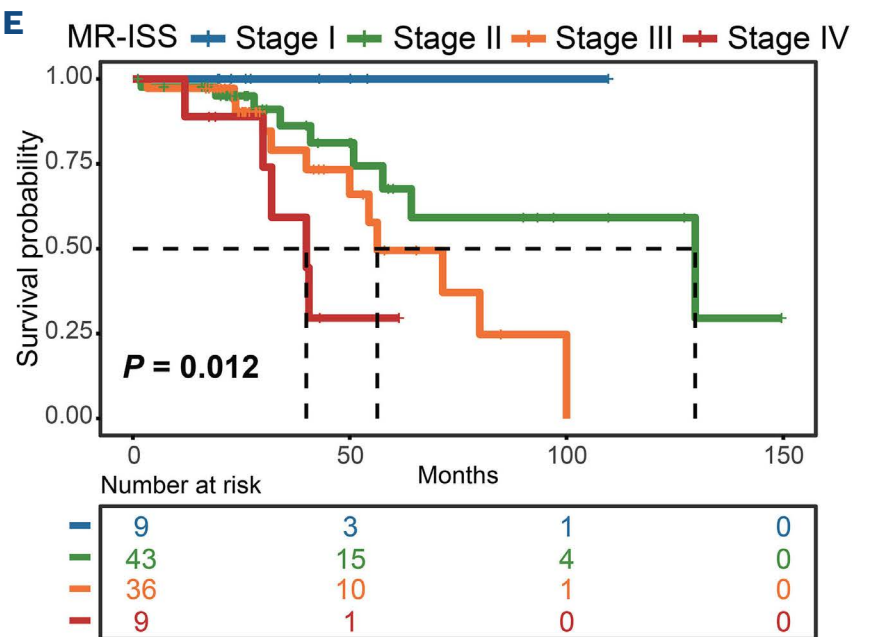
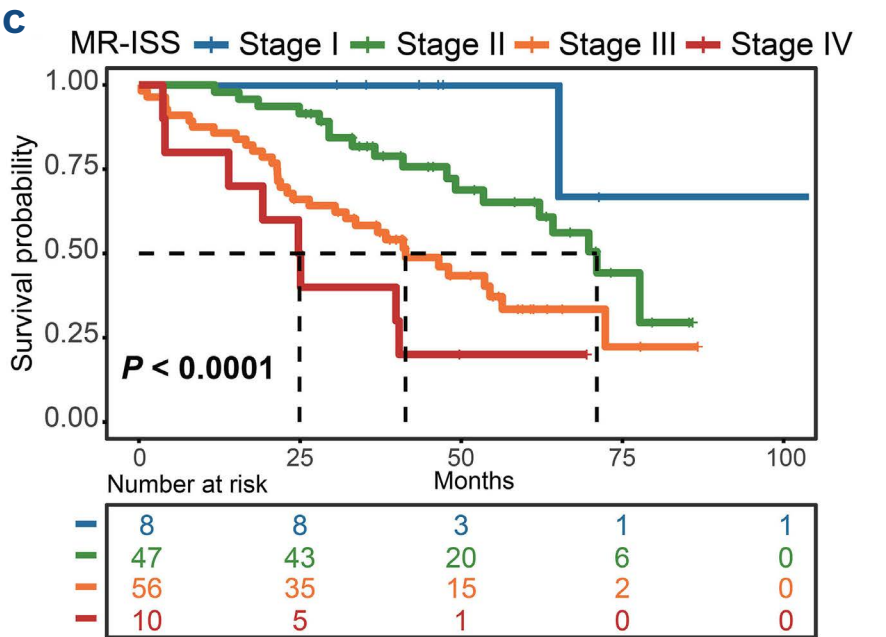


Figure 7. Modified Revised International Scoring System model from sub-C4 plasma cells predicting short survival in myeloma patients. (A) Kaplan-Meier curve showing the overall survival of multiple myeloma (MM) patients in the MMRF-CoMMpass dataset by modified Revised International Scoring System (MR-ISS) groups. (B) Kaplan-Meier curve showing the overall survival of MM patients in the validation set 1 (GSE136337 cohort) by MR-ISS groups. (C) Kaplan-Meier curve showing the overall survival of MM patients in the validation set 2 (in-house data) by MR-ISS groups. (D, E) The detection and prognostic significance of the seven genes were confirmed using the droplet digital polymerase chain reaction (ddPCR) assay. Monochrome (D) and multicolor (E) ddPCR detection was used for quantitative assessment of seven-gene scores in 139 newly diagnosed patients. Kaplan-Meier curves depict the overall survival rate of the MM patients.



plays a critical role in tumor cell proliferation and the immunosuppressive tumor microenvironment.^{10,22,23} Here, our study further revealed that LILRB4 is the only member of the LILRB family that is highly expressed on MM cells and its expression correlates with the outcome of myeloma patients. LILRB4 promotes MM cell proliferation by activation of the downstream NF- κ B signaling pathway, which is also involved in proteasome inhibitor-resistance of MM cells. Although there has been much progress in identifying LILRB4 function in myelomagenesis, the molecular mechanisms have not been fully understood. LILRB4 is a special member of the LILRB family, which contains two extracellular immunoglobulin domains, a transmembrane domain, and three cytoplasmic immunoreceptor tyrosine-based inhibitory motifs. The functional ligand of LILRB4 has not been discovered. We found that APOE, one of the candidate ligands³² of LILRB4, had limited effects on triggering the activation of LILRB4 and the downstream pathway in MM cells. Therefore, further study needs to be done to clarify the molecular mechanisms of LILRB4 in myeloma.

Besides the effect on promoting MM cell survival, recent studies discovered that LILRB4 also contributes to bone damage by promoting the differentiation and maturation of osteoclasts.^{22,33} Consistently, our study revealed that patients in the EM24 group with high levels of LILRB4 exhibited more severe bone disease compared to that of patients in the nEM24 group (*data not shown*). Currently, LILRB4 has emerged as a hot therapeutic target for tumors and autoimmune diseases, leading to the initiation of clinical trials for several anti-LILRB4 monoclonal antibodies and chimeric antigen receptor T-cell immunotherapy.^{34,35} Besides LILRB4, our recent study revealed that CRIP1 is involved in the proteostasis of MM cells through dual regulation of the proteasome and autophagy, which promotes tumor cell growth, against proteasome inhibitor treatment and facilitates development of metastasis.³¹ Collectively, these findings by us and other groups indicate that the seven genes identified in EM24 patients are involved in the

aggressive behavior of MM, not only by directly promoting malignant proliferation of MM cells and mediating drug resistance, but also, more importantly, their high expression is closely associated with immune cell dysfunction in the tumor microenvironment.

In summary, this study sheds light on the unique aspects of heterogeneity in MM cells and the altered functional states of immune cells. We have constructed a novel risk stratification tool, the MR-ISS score, which can efficiently identify ultra-high-risk patients with very poor outcome, which might guide the development of novel therapeutic strategies to benefit these ultra-high-risk patients.

Disclosures

No conflicts of interest to disclose.

Contributions

LL, HS and MH analyzed data and drafted the paper. FF, XS, JM, RL and TY collected data and followed up the patients. HS, XZ, HJ, YY and FM collected data and completed the statistical analysis from the validation cohort. LY, XL and ZY designed the ddPCR. LQ and MH designed the research, and revised the manuscript.

Funding

This work was supported by the following foundations: the Noncommunicable Chronic Diseases-National Science and Technology Major Project (2023ZD0501300), the Natural Science Foundation of China (82341211, 82370210, 82170194, 82170203), and the CAMS Innovation Fund for Medical Sciences (2021-I2M-1-040, 2023-I2M-2-007, 2022-I2M-1-022).

Data-sharing statement

The in-house data are deposited in the Genome Sequence Archive repository (<https://bigd.big.ac.cn/gsa-human/browse/HRA003504>), BioProject ID: PRJCA013382, accession ID: HRA003504.

References

1. Kaiser MF, Hall A, Walker K, et al. Daratumumab, cyclophosphamide, bortezomib, lenalidomide, and dexamethasone as induction and extended consolidation improves outcome in ultra-high-risk multiple myeloma. *J Clin Oncol.* 2023;41(23):3945-3955.
2. Rajkumar SV. Multiple myeloma: 2024 update on diagnosis, risk-stratification, and management. *Am J Hematol.* 2024;99(9):1802-1824.
3. Sonneveld P, Avet-Loiseau H, Lonial S, et al. Treatment of multiple myeloma with high-risk cytogenetics: a consensus of the International Myeloma Working Group. *Blood.* 2016;127(24):2955-2962.
4. Rees MJ, Kumar S. High-risk multiple myeloma: redefining genetic, clinical, and functional high-risk disease in the era of molecular medicine and immunotherapy. *Am J Hematol.* 2024;99(8):1560-1575.
5. Hagen P, Zhang J, Barton K. High-risk disease in newly diagnosed multiple myeloma: beyond the R-ISS and IMWG definitions. *Blood Cancer J.* 2022;12(5):83.
6. Palumbo A, Avet-Loiseau H, Oliva S, et al. Revised International Staging System for multiple myeloma: a report from International Myeloma Working Group. *J Clin Oncol.* 2015;33(26):2863-2869.
7. D'Agostino M, Cairns DA, Lahuerta JJ, et al. Second Revision of the International Staging System (R2-ISS) for overall survival in multiple myeloma: a European Myeloma Network (EMN) report within the HARMONY project. *J Clin Oncol.* 2022;40(29):3406-3418.
8. Shaughnessy JD JR, Zhan F, Burington BE, et al. A validated gene expression model of high-risk multiple myeloma is defined

- by deregulated expression of genes mapping to chromosome 1. *Blood*. 2007;109(6):2276-2284.
9. Kuiper R, Broyl A, de Knecht Y, et al. A gene expression signature for high-risk multiple myeloma. *Leukemia*. 2012;26(11):2406-2413.
 10. Gong L, Sun H, Liu L, et al. LILRB4 represents a promising target for immunotherapy by dual targeting tumor cells and myeloid-derived suppressive cells in multiple myeloma. *Haematologica*. 2024;109(11):3650-3669.
 11. Lv J, Sun H, Gong L, et al. Aberrant metabolic processes promote the immunosuppressive microenvironment in multiple myeloma. *Front Immunol*. 2022;13:1077768.
 12. Sun H, Fang T, Wang T, et al. Single-cell profiles reveal tumor cell heterogeneity and immunosuppressive microenvironment in Waldenström macroglobulinemia. *J Transl Med*. 2022;20(1):576.
 13. Liu L, Gong D, Sun H, et al. DNp73 enhances tumor progression and immune evasion in multiple myeloma by targeting the MYC and MYCN pathways. *Front Immunol*. 2024;15:1470328.
 14. Yu T, Du C, Ma X, et al. Polycomb-like protein 3 induces proliferation and drug resistance in multiple myeloma and is regulated by miRNA-15a. *Mol Cancer Res*. 2020;18(7):1063-1073.
 15. Sudha P, Ahsan A, Ashby C, et al. Myeloma Genome Project Panel is a comprehensive targeted genomics panel for molecular profiling of patients with multiple myeloma. *Clin Cancer Res*. 2022;28(13):2854-2864.
 16. Palumbo A, Anderson K. Multiple myeloma. *N Engl J Med*. 2011;364(11):1046-1060.
 17. Greipp PR, San Miguel J, Durie BG, et al. International Staging System for multiple myeloma. *J Clin Oncol*. 2005;23(15):3412-3420.
 18. Emde-Rajaratnam M, Beck S, Benes V, et al. RNA-sequencing based first choice of treatment and determination of risk in multiple myeloma. *Front Immunol*. 2023;14:1286700.
 19. Biran N, Dhakal B, Niesvizky R, et al. Enhancing risk stratification and treatment decision in multiple myeloma with SKY92 gene expression profiling in real-world data. *Br J Haematol*. 2025;206(6):1642-1653.
 20. Al Hadidi S, Ababneh OE, Schinke CD, et al. Three years of maintenance with VRD in multiple myeloma: results of total therapy IIIB with a 15-year follow-up. *Blood Adv*. 2024;8(3):703-707.
 21. Cerchione C, Usmani SZ, Stewart AK, et al. Gene expression profiling in multiple myeloma: redefining the paradigm of risk-adapted treatment. *Front Oncol*. 2022;12:820768.
 22. Wang H, Wang L, Luan H, et al. LILRB4 on multiple myeloma cells promotes bone lesion by p-SHP2/NF- κ B/RELt signal pathway. *J Exp Clin Cancer Res*. 2024;43(1):183.
 23. Xie L, Chen C, Zhang T, et al. LILRB4 regulates multiple myeloma development through STAT3-PFKFB1 pathway. *Cell Death Dis*. 2024;15(7):515.
 24. Li J, Yang Y, Wang W, et al. Single-cell atlas of the immune microenvironment reveals macrophage reprogramming and the potential dual macrophage-targeted strategy in multiple myeloma. *Br J Haematol*. 2023;201(5):917-934.
 25. Zhang X, Han Y, Fan C, Jiang Y, Jiang W, Zheng C. Epigallocatechin gallate induces apoptosis in multiple myeloma cells through endoplasmic reticulum stress induction and cytoskeletal disruption. *Int Immunopharmacol*. 2024;141:112950.
 26. Brito JL, Walker B, Jenner M, et al. MMSET deregulation affects cell cycle progression and adhesion regulons in t(4;14) myeloma plasma cells. *Haematologica*. 2009;94(1):78-86.
 27. Boiarsky R, Haradhvala NJ, Alberge JB, et al. Single cell characterization of myeloma and its precursor conditions reveals transcriptional signatures of early tumorigenesis. *Nat Commun*. 2022;13(1):7040.
 28. Li H, Guo L, Li B, Li X. SUMO1 modification of histone H4 is involved in the pathogenesis of nodular lymphocyte predominant Hodgkin lymphoma. *Transl Cancer Res*. 2020;9(7):4413-4423.
 29. Neri P, Ren L, Azab AK, et al. Integrin β 7-mediated regulation of multiple myeloma cell adhesion, migration, and invasion. *Blood*. 2011;117(23):6202-6213.
 30. Ohguchi H, Hideshima T, Bhasin MK, et al. The KDM3A-KLF2-IRF4 axis maintains myeloma cell survival. *Nat Commun*. 2016;7:10258.
 31. Tang P, Yu Z, Sun H, et al. CRIP1 involves the pathogenesis of multiple myeloma via dual-regulation of proteasome and autophagy. *EBioMedicine*. 2024;100:104961.
 32. Gui X, Deng M, Song H, et al. Disrupting LILRB4/APOE Interaction by an efficacious humanized antibody reverses T-cell suppression and blocks AML development. *Cancer Immunol Res*. 2019;7(8):1244-1257.
 33. Deng M, Gui X, Kim J, et al. LILRB4 signalling in leukaemia cells mediates T cell suppression and tumour infiltration. *Nature*. 2018;562(7728):605-609.
 34. Morse JW, Rios M, Ye J, et al. Antibody therapies for the treatment of acute myeloid leukemia: exploring current and emerging therapeutic targets. *Expert Opin Investig Drugs*. 2023;32(2):107-125.
 35. Sharma N, Atolagbe OT, Ge Z, Allison JP. LILRB4 suppresses immunity in solid tumors and is a potential target for immunotherapy. *J Exp Med*. 2021;218(7):e20201811.