LILRB4 represents a promising target for immunotherapy by dual targeting tumor cells and myeloid-derived suppressive cells in multiple myeloma

Lixin Gong,[1,2*] Hao Sun,[1,2*] Lanting Liu,[1,2] Xiyue Sun,[1,2] Teng Fang,[1,2] Zhen Yu,[1,2] Weiwei Sui,[1,2] Jingyu Xu,[1,2] Tingyu Wang,[1,2] Fangshuo Feng,[1,2] Lei Lei,[3] Wei Rui,[3] Yuxuan Liu,[1,2] Xueqiang Zhao,[4] Gang An,[1,2] Xin Lin,[4] Lugui Qiu[1,2] and Mu Hao[1,2]

[1]State 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; [2]Tianjin Institutes of Health Science, Tianjin; [3]BriSTAR Immunotech Biotechnology Co. Ltd., Beijing and [4]Department of Basic Medical Sciences, Tsinghua University School of Medicine, Beijing, China

**LG and HS contributed equally as first authors.*

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

L. Qiu
qiulg@ihcams.ac.cn

X. Lin
linxin307@tsinghua.edu.cn

Received: January 20, 2024.
Accepted: May 20, 2024.
Early view: May 30, 2024.

https://doi.org/10.3324/haematol.2024.285099

Supplemental files

Methods

Sample collection and single cell preparation

BMMCs were isolated by Ficoll (MERCK) density-gradient centrifugation and cryopreserved at -80°C for less than five days until processed. The number and viability of cells was measured using a TC20 automated cell counter (Biorad). Dead cells (cell viability less than 80%) were removed by magnetic bead purification (Miltenyi Biotech) according to the manufacturer's protocol before scRNA-seq.

Single-cell RNA library preparation and sequencing

Chromium single-cell sequencing technology was performed following the manufacturer's protocol (10× Genomics). Library construction procedures were performed using the Chromium Single Cell 3' Library, Gel Bead & Multiplex Kit (10× Genomics, V2), strictly following the manufacturer's instructions. Then the cell suspensions were loaded onto the 10x Chromium Single Cell Controller to generate single-cell gel bead-in-emulsions (GEMs), and we performed barcoded reverse transcription of RNA within a single cell using a Verity Thermal Cycler (Life Technologies). Through reverse transcription in a single GEM, the barcodes were added to the RNAs released from lysed cells; then fragmentation, end repair, polyA tailing, and adaptor ligation were achieved according to the standard protocol. The cDNA purification and size selection were performed by SPRI select beads (Beckman Coulter), and the quality was evaluated using the Agilent Bioanalyzer. Finally, the libraries were sequenced on an MGISEQ-2000 sequencer as 150 bp paired-end reads by Beijing Genomics Institute (BGI, Shenzhen, China).

scRNA-seq data processing

The Cell Ranger Software Suite (version 3.0.2; 10x Genomics) was used to perform sample de-multiplexing, alignment, barcode processing, and unique molecular identifier (UMI) counting. Briefly, sequencing reads were aligned against the GRCh38 human reference genome with STAR, and count matrices were built from the resulting BAM files.(1) The R package scCancer v2.1.0(2) was then used to conduct the quality control, which detected outliers automatically according to the distributions of the metrics. First, the number of total unique molecular identifiers (UMI; nUMI), number of expressed genes (nGene), and percentage of UMIs from mitochondrial genes (mito.percent), ribosomal genes (ribo.percent) and dissociation-associated genes (diss.percent) were used to filter low-quality cells and multiplets (see the auto-selected thresholds in **Suppl. Fig. 1B**). We then filtered the genes that were expressed in <3 cells. For the integration of the cells from different samples, we corrected batch effect using the R package

harmony (version 0.1.1).(3) In parameter settings, the first 30 dimensions of canonical correlation analysis (CCA) and principal component analysis (PCA) were used.

Dimensionality reduction, clustering of cells, and visualization

The R package Seurat v.3 was used for data scaling, transformation, clustering, dimensionality reduction, differential expression analysis, and most visualizations.(4, 5) The variable genes were identified using the 'vst' method in the Seurat 'FindVariableFeatures' function. PCA was performed using the top 2,000 variable genes. Graph-based clustering was performed on the PCA-reduced data for clustering analysis. The resolution was set to 0.5 to obtain a more refined result. Briefly, the first 50 PCs of the integrated gene-cell matrix were used to construct a shared nearest-neighbor graph (SNN; 'FindNeighbors'), and this SNN was used to cluster the dataset ('FindClusters') using a graph-based modularity-optimization algorithm of the Louvain method for community detection. Then UMAP was performed on the top 30 principal components for visualizing the cells.

Cell cluster annotation with specific maker genes expression

The marker genes, identified by the 'FindAllMarkers' function—with the setting 'only.pos' as 'True', 'min.pct' as 0.25, 'logfc.threshold' as 0.25 and 'test.use' as 'wilcox'—were used to annotate cell clusters. Cluster annotation was confirmed using the celltypist, an automated cell type annotation tool for scRNA-seq datasets based on logistic regression classifiers optimised by the stochastic gradient descent algorithm.(6)

Characterizing the cell distribution in sample groups

To characterize the group distribution of cells, odds ratios (OR) were calculated and used to indicate preferences as reported.(7) Specifically, for each combination of cells i and group j, a 2 by 2 contingency table was constructed, which contained the number of cells i in group j, the number of cells of cell i in other groups, the number of non-i cells in group j, the number of non-i cells in other groups. Then Fisher's exact test was applied on this contingency table, thus OR and corresponding p-value could be obtained. We labeled groups with a p-value < 0.05 using an asterisk (*). A higher OR with an asterisk indicated that cells i was more preferred to distribute in group j, while a lower OR with an asterisk indicated that cells i was preferred to less distribute in group j.

Plasma-cell lineage developmental trajectory

The scRNA-seq data of $CD34^+$, $CD19^+$, and $CD138^+$ cells from healthy donors including PRJCA003794(8) and PRJNA732205(9) were used to build the plasma-cell lineage and differentiation stage classifier. Pseudotime-ordered analysis of HSCs, pro-B cells, pre-B cells, immature B cells, naïve B cells, memory

B cells, plasmablasts, and plasma cells was performed using Monocle2.(10)

Plasma-cell differentiation stage classifier

To discriminate the differentiation stage of MM cell clusters, we trained a plasma-cell differentiation stage classifier using random forest training. The training was based on eight annotated cells representing different differentiation stages (HSCs, pro-B cells, pre-B cells, immature B cells, naïve B cells, memory B cells, plasmablasts, and plasma cells). We utilized the 'FindAllMarkers' function from Seurat v.3 to select the top 100 marker genes for each stage, sorted according to the average fold change. Cell cycle-related genes, mitochondrial genes, ribosomal genes, and genes associated with cell dissolution were removed, resulting in 597 non-redundant feature genes in total. The normalized gene-cell matrix by Seurat v.3 was randomly split into a 70% training set and a 30% validation set using the caTools package. The training was performed using the 'randomForest' function from the randomForest v4.6 package, generating 500 trees. For testing, we used the 'roc' function from the pROC package v1.18.0 to calculate the sensitivity and specificity of the predictions for each cell type in the validation set. Finally, we utilized ggroc v1.0 to plot the ROC curves.

Processing of scBCR-seq data

The outsourced scBCR-seq were processed using CellRanger v3.0.2 (‘vdj’ pipeline with default setting and using the V(D)J library GRCh38) and the output files with cell and clonotype information were generated. According to the clonotypes provided by the file ‘filtered_contig_annotations.csv’, we counted the cell number of each clonotype as its abundance. Cells were assigned to ‘abundance>10′, ‘abundance≤10’ or ‘none’ according to the abundance of their clonotypes.

Calculating the proportion of Kappa$^+$ cells

The human genome has one kappa constant (IGKC) gene but variable number of lambda constants – IGLC1, IGLC2, IGLC3 and IGLC7 are functional isotypes. Let *t* be a chosen transcript count threshold and *K* the number of cells expressing at least t IGKC transcripts. Further let *L* be the number of cells expressing at least t transcripts of any of three out of the four functional lambda isotypes: IGLC2, IGLC3 and IGLC7.(11) The IGLC1 transcript was not detected in any B cell, probably because complete overlap of its 3’end with the IGLL5 gene precludes IGLC1 mRNA quantification with 3’end RNA-Seq assays. The % kappa$^+$ cells within a given population was calculated as: $Kappa^+ = \frac{K}{K+L}$. t = 1 was chosen. The estimated % of kappa$^+$ cells was robust to different values of *t* ranging between 1 and 5.

Topology of sequenced genes in MM cells

We transposed the gene-by-cell matrix of MM cells and applied dimensionality reduction and clustering techniques to analyze 20,711 genes across 12,835 MM cells using Seurat v.3. To identify genes with differential expression in specific clusters or samples, we employed the 'FindMarkers' function with the 'wilcox' parameter, conducting the Wilcoxon rank-sum test with a significance threshold of $P<0.05$. Differential genes with a fold change greater than 1 were considered specific to clusters or samples. We utilized UMAP for visualizing their distribution characteristics.

Construction of transcriptional regulatory network and gene co-expression network

We selected specific genes from cluster 4 to infer the transcriptional regulatory network and establish a co-expression network. Candidate transcription factors and transcription factor-target pairs were obtained from TRRUST v2(12), which includes 8,444 regulatory interactions (annotated as ‘activation’, ‘repression’, and ‘unknown’) for 800 transcription factors in humans. We retained only the edges where both the transcription factor and target were present in the specific genes of cluster 4. Moreover, we kept only the transcription factors with at least two edges annotated as ‘activation’ or ‘repression’ in the final network.

To construct the co-expression network, we calculated the Pearson correlation between the specific genes of cluster 4 in cluster 4 MM cells. Genes with a P-value < 0.05 and an absolute correlation coefficient > 0.4 were retained. Clustering of the co-expression network was performed using MCODE in Cytoscape v3.9.0, and the top 4 clusters with a degree cutoff of 5 were kept. The networks were visualized using Gephi v0.10.1 with the ‘Fruchterman Reingold’ layout.

WES data processing and mutation calling from scRNA-seq data

Mutations of reported driver genes(13) were identified from bulk Whole Exome Sequencing (WES) data from 947 newly diagnosed MM patients of MMRF-CoMMpass cohort processed by MuTect2 of GATK pipeline (vcf files of IA20 version, https://tcga-data.nci.nih.gov/). Next, the identified variants were further examined in the scRNA-seq data of the MM cells using cb_sniffer v1.0 (https://github.com/sridnona/cb_sniffer). For this analysis, two additional files generated by CellRanger were required: the aligned BAM file of scRNA-seq data and the corresponding cell barcode file. The output file of cb sniffer with the suffix name ‘counts_CB.tsv’ recorded cell barcodes that had the specific mutation. Finally, the frequency of 63 driver genes mutation in single cell was calculated as mutation score.

The significance of cluster 4 on overall survival of MM patients

Deconvolution of subclusters in bulk tumor cells from GSE2658 RNA martrix was performed using support

vector machine (SVM) algorithms by CIBERSORTx(14). Normalized average expression matrix of 10 plasma cell clusters served as the input signature matrix and the relative cell fraction was further imputed with scRNA-seq mode. The optimal cut point for the fraction of cluster 4 in MM patients was determined using the maximally selected rank statistics from the maxstat R package. The log-rank test was employed to calculate *P*-values between MM patients with high or low cluster 4 fractions, and the Kaplan–Meier method was performed to plot survival curves for survival analyses.

Single cell copy number variations (CNVs) calling

To identify clonal large-scale chromosomal CNVs in malignant MM cells, we used the inferCNV R package to infer the genetic profiles of each cell based on the average expression of large genes sets (101 genes) in each chromosomal region of the tumor genome compared to normal cells(15) All malignant plasma cells MM were input as interrogation group and normal plasma cells from HDs were sampled as control. Other parameters were set as default. CNVs on autosomes were visualized using the R package RCircos(16).

Projection and analysis of myeloid cells

We leveraged STACAS package to integrate pan-cancer myeloid cells from 45 human samples of PRJNA647394 and myeloid-derived suppressor cells (MDSCs) from PRJNA578550 and further built a customized reference atlas for myeloid cells.(17-19) We projected our myeloid cell to this reference map with ProjecTILs v3.1, allowing accurate embedding of new scRNA-seq data into a reference.(20) Subsequently, we calculated MDSC signature, activated score and migratory score of DC cells, anti-/pro-inflammatory score of monocytes, and M1/M2 polarization score of macrophages with reported gene set by AddModuleScore function in Seurat v.3.(18, 19, 21)

RNA sequencing and data processing

Bulk RNA-seq was used to characterize the in-house data of the MM patients for verification. 133 newly diagnosed MM patients were involved at the Institute of Hematology and Blood Diseases Hospital, Chinese Academy of Medical Sciences, and Peking Union Medical College. RNA extraction was performed using the miRNeasy Mini Kit (Qiagen). The RNA yield and cDNA libraries from CD138-selected plasma cells were sequenced on the NextSeq500 Sequencing System (Illumina) with read length of paired-end 150 bp. Raw reads were pre-processed using FastQC software. Clean reads were then used for subsequent analyses. RNA-seq reads were aligned to GRCh38 using STAR. Transcript expression levels were quantified after normalizing the count data with the edgeR package.(22)

ATAC-seq analysis

Assay for Transposase Accessible Chromatin with high-throughput sequencing (ATAC-seq) uses Tn5 transposase to insert adaptors into the accessible region on the genome. The obtained library can directly identify the transcription factor binding region and nucleosome position, providing an effective approach for researching gene regulation and genomic imprinting. ATAC-seq unveiled significant increase in chromatin accessibility within the promoter region of the LILRB4 gene in cell lines exhibiting elevated expression levels of LILRB4.

Cell Culture and transfection

MM cell lines were kept in our laboratory and authenticated by short-tandem repeat (STR) profiling. The MM cell lines with stable LILRB4 overexpression (OE) or deletion (KO) were produced by lentivirus transfection containing LILRB4-OE plasmid (pCDH-EF1A-h*LILRB4*-T2A-Puro) or Cas9-expressing plasmids containing guide RNA (LentiCRISPR v2) respectively. The guide RNA sequences used for LILRB4 were listed in **Table S4**. Lentivirus packaging, the constructs for gene overexpression and CRISPR-Cas9-based constructs were performed according to our previous study.(23, 24) For LILRB4 knock-out, LILRB4 or control CRISPR lentivirus were transduced into cells and sorted via FACS isolation. Single clones were selected and validated by flow detection.

Colony-formation assay

A total of 1 000 MM cells were plated in triplicate in 1.1 ml methylcellulose-based medium (MethoCult™ H4230, StemCell Technologies) per 12-well and incubated for 2 weeks. The culture plate was taken photo under high-content analysis system (Operetta CLS). Colonies consisting of more than 50 cells were scored.

Co-culture assay

For MDSC induction, MM cells (1×10^5) and PBMCs from healthy donors (1×10^6) were co-cultured in 10% RPMI 1640 for 6 days. MDSC detection ($CD11b^+CD33^+HLA\text{-}DR^{-/low}$) was performed by flow cytometry analysis. The in vitro experiments were performed in triplicate and repeated three times.

For examining the regulation of LILRB4 expression by microenvironment cells, MM cells were firstly labeled by CFSE. Then MM cells and PBMCs from healthy donors were co-cultured at the ratio of 1:10. After 3-day coculture, LILRB4 expression in $CFSE^+$ MM cells was examined by flow cytometry. At the same time, $CFSE^+$ MM cells were sorted by flow cytometry for *LILRB4* mRNA detection.

MDSC generation

For MDSC cytotoxicity assay, MDSCs were induced in vitro. $CD14^+$ monocytes were isolated from HD PBMCs following the EasySep Human Monocyte Isolation Kit (Stem cell, 19359). After isolation, monocytes were cultured in RPMI 1640 complete medium with the addition of IL-6 (10ng/ml) & GM-CSF (10ng/ml) & PGE2 (1μg/ml) for 7 days. Flow sorting of MDSCs were then performed according to the expression of $CD14^+CD33^+CD11b^+$.

Cytotoxicity assays

Target cells were labelled with CellTrace FarRed dye and incubated with LILRB4 STAR-T cells or mock-T cells at the indicated effector: target (E:T) ratio in triplicates for 24h or 48h. At the end of the co-culture, cells were harvested and stained with propidium iodide (PI). PI^+CTR^+ cells were quantified by flow cytometer. Co-culture supernatants were measured for the concentration of IFN-γ, IL-2, TNF-α following the enzyme-linked immunosorbent assay kits.

To determine cytotoxicity activity of LILRB4 STAR-T cells against MM patient samples, BMMCs were first stained with antibodies against CD38, CD11b, CD14, LILRB4 to quantify the LILRB4 positive target cells in BMMCs. Then, BMMCs were co-cultured with STAR-T cells or mock-T cells for 4h at E:T ratio of 1:1 and 5:1. After incubation, the co-culture cells were stained with CD38, CD11b, CD14, LILRB4 and 7-AAD. Counting beads were added for normalization as well.

Flow cytometry analysis

Flow cytometry was performed on CantoⅡ flow cytometer (BD Biosciences) and the data were analyzed by Flowjo V10 software (Treestar). The antibodies are listed in **Table S5**.

Statistical analysis

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

Supplemental Tables

Table. S1| The gene list of the highly expressed genes regulated by H3K27ac in cluster4

Gene	p_val	avg_logFC	avg_logExp.1	avg_logExp.2	p_val_adj	geneChr
LILRB4	0	2.17908523	2.40554514	0.22645992	0	chr19
STMN1	0	1.40553826	2.75145486	1.34591659	0	chr1
PTTG1	0	0.95213495	1.28386214	0.33172719	0	chr5
PAGE5	0	0.8527573	0.89321904	0.04046175	0	chrX
TK1	0	0.72126706	0.88658346	0.1653164	0	chr17
MCM5	0	0.65412165	0.84310093	0.18897929	0	chr22
CENPW	0	0.62891533	0.74713037	0.11821504	0	chr6
PHF19	0	0.57845067	0.69203694	0.11358627	0	chr9
PVALB	0	0.53796586	0.55888082	0.02091495	0	chr22
GINS2	0	0.53282238	0.63076266	0.09794028	0	chr16
SSX1	0	0.290379	0.30059916	0.01022016	0	chrX
MAGEB2	0	0.28337759	0.29427583	0.01089824	0	chrX
CENPH	2.74E-302	0.55582099	0.67808887	0.12226788	5.76E-298	chr5
PAFAH1B3	3.44E-301	0.80709162	0.95451379	0.14742217	7.25E-297	chr19
GAPDH	4.39E-301	2.02869199	4.72114847	2.69245649	9.25E-297	chr12
CDT1	5.20E-299	0.51638582	0.61495447	0.09856865	1.10E-294	chr16
CENPN	4.99E-296	0.56668702	0.69338616	0.12669914	1.05E-291	chr16
CRIP1	9.97E-295	2.11911896	3.14333347	1.0242145	2.10E-290	chr14
GULP1	1.84E-289	0.32850986	0.34561042	0.01710056	3.87E-285	chr2
SHCBP1	1.90E-282	0.38314151	0.43386619	0.05072469	4.00E-278	chr16
ENO1	2.74E-282	1.61850607	2.49969514	0.88118907	5.77E-278	chr1
FAM3B	2.75E-282	0.31034582	0.31738059	0.00703477	5.79E-278	chr21
ACY3	5.01E-279	0.66839987	0.76295517	0.0945553	1.05E-274	chr11
RRM2	1.29E-277	0.87093479	1.08603547	0.21510068	2.71E-273	chr2
MAD2L1	1.96E-269	0.62099465	0.82557036	0.2045757	4.13E-265	chr4
TYMS	5.49E-268	0.79456845	1.16973305	0.37516459	1.15E-263	chr18
CDC6	5.83E-265	0.39851254	0.46916167	0.07064913	1.23E-260	chr17
C19orf48	2.80E-260	0.91900105	1.2763505	0.35734945	5.89E-256	chr19
ASF1B	2.31E-259	0.41431229	0.50840515	0.09409286	4.86E-255	chr19
GMNN	6.86E-257	0.56085582	0.7075087	0.14665288	1.44E-252	chr6
CD9	1.90E-254	1.38218422	1.8881117	0.50592748	4.00E-250	chr12
TUBA1B	1.72E-253	1.72586739	3.58602697	1.86015958	3.63E-249	chr12
PCLAF	1.36E-250	1.11541067	1.72628242	0.61087175	2.85E-246	chr15
MCM7	2.34E-250	0.72937413	1.08684466	0.35747054	4.93E-246	chr7
C11orf96	7.10E-249	0.72561813	0.82765119	0.10203306	1.50E-244	chr11
ADTRP	8.92E-249	0.53174978	0.61626589	0.08451611	1.88E-244	chr6
NLRP7	3.51E-247	0.2750494	0.28101606	0.00596667	7.38E-243	chr19
POLD2	2.31E-245	0.7992543	1.0874332	0.2881789	4.87E-241	chr7
FBL	5.27E-241	1.11599251	1.62947462	0.5134821	1.11E-236	chr19

Table. S2| H3K27ac regulation in LILRB4 expression from ChIP-seq data of GSE145891

Dataset	PeakScore (MACS 2)	-log10 (Pvalue)	PeakStart	PeakEnd	annotation	distanceToTSS	GENESYMBOL
GSM4338508	42	6.23028	55171781	55178336	Promoter (<=1kb)	0	LILRB4
GSM4338510	17	4.41754	55172220	55172804	Promoter (1-2kb)	-1320	LILRB4
GSM4338514	31	5.34083	55172181	55173612	Promoter (<=1kb)	-512	LILRB4
GSM4338516	63	8.31412	55171391	55178385	Promoter (<=1kb)	0	LILRB4
GSM4338518	30	5.61436	55173013	55174228	Promoter (<=1kb)	0	LILRB4
GSM4338520	12	2.90607	55172345	55172891	Promoter (1-2kb)	-1233	LILRB4
GSM4338522	18	3.64749	55172593	55172801	Promoter (1-2kb)	-1323	LILRB4

Table. S3| The baseline clinical information of bone marrow samples used for LILRB4 STAR-T cell cytotoxicity analysis

Sample	Disease status	Treatment	Cytogenetic aberrancy
Pt1	Post-treatment	PI-based	1q21gain, del(13q)
Pt2	Post-treatment	PI-based	1q21gain, del(13q)
Pt3	ND PCL	n.a.	t(11;14)
Pt4	NDMM	n.a	del(13q)
Pt5	RRMM	PI-based	1q21gain, del(13q), t(14;unidentified)
Pt6	RRMM	PI-based	1q21gain, del(13q), t(14;unidentified)
Pt7	RRMM	PI-based, Daratumumab, BCMA-ADC	1q21gain, del(13q), t(14;unidentified)
Pt8	RRMM	PI-based, BCMA CAR-T	1q21gain, t(14;unidentified)
Pt9	RRMM	PI-based, BCMA CAR-T	1q21gain, t(4;14)

Table. S4 | The guiding RNA for LILRB4 Crispr-Cas9 knockout

Guiding RNA #	Guiding RNA sequence
sgRNA1	5′-TGTTACTATCGCAGCCCTGT-3′
sgRNA2	5′-GTAGGTCCCCCCGTGCACTG-3′
sgRNA3	5′-CCTGTGACCTCAGTGCAC GG-3‘

Table. S5 | Antibodies for flowcytometry analysis

Antibody	Clone	Company	Order number
anti-hCD19-BV421	UCHT1	BioLegend	300452
anti-hCD56-FITC	TULY56	eBioscience	4299203
anti-hCD38-PE-Cy7	HIT2	BioLegend	303516
anti-hCD138-PE	MI15	BD Biosciences	552026
anti-hCD45-Percp	2D1	BD Biosciences	347464
anti-hLILRB4-APC	ZM4.1	BioLegend	333016
anti-hCD11b-PE-Cy7	M1/70	BioLegend	101216
anti-hCD14-APC-Cy7	HCD14	BioLegend	325619
anti-hCD15-FITC	MMA	BioLegend	394706
anti-hCD33-PE	WM53	BD Biosciences	555450
anti-HLA-DR-Pacific-blue	LN3	BioLegend	327016
anti-mCD45-PE-Cy7	30-F11	BioLegend	103114
anti-mCD11b-APC	M1/70	eBioscience	17-0112-82
anti-mGr1-PE	RB6-8C5	BioLegend	108407
Live/Dead Cell Stain Kit		ThermoFisher Scientific	L34965

References

1. Dobin A, Davis CA, Schlesinger F et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics (Oxford, England). 2013 Jan 1;29(1):15-21.

2. Guo W, Wang D, Wang S et al. scCancer: a package for automated processing of single-cell RNA-seq data in cancer. Brief Bioinform. 2021 May 20;22(3).

3. Korsunsky I, Millard N, Fan J et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods. 2019 Dec;16(12):1289-1296.

4. Butler A, Hoffman P, Smibert P et al. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018 Jun;36(5):411-420.

5. Stuart T, Butler A, Hoffman P et al. Comprehensive Integration of Single-Cell Data. Cell. 2019 Jun 13;177(7):1888-1902.e21.

6. Dominguez Conde C, Xu C, Jarvis LB et al. Cross-tissue immune cell analysis reveals tissue-specific features in humans. Science. 2022 May 13;376(6594):eabl5197.

7. Zheng L, Qin S, Si W et al. Pan-cancer single-cell landscape of tumor-infiltrating T cells. Science. 2021 Dec 17;374(6574):abe6474.

8. Zhang Y, Wang S, Zhang J et al. Elucidating minimal residual disease of paediatric B-cell acute lymphoblastic leukaemia by single-cell analysis. Nat Cell Biol. 2022 Feb;24(2):242-252.

9. Alameda D, Goicoechea I, Vicari M et al. Tumor cells in light-chain amyloidosis and myeloma show distinct transcriptional rewiring of normal plasma cell development. Blood. 2021 Oct 28;138(17):1583-1589.

10. Qiu X, Mao Q, Tang Y et al. Reversed graph embedding resolves complex single-cell trajectories. Nature methods. 2017 Oct;14(10):979-982.
11. Rai A, Greening DW, Chen M et al. Exosomes Derived from Human Primary and Metastatic Colorectal Cancer Cells Contribute to Functional Heterogeneity of Activated Fibroblasts by Reprogramming Their Proteome. Proteomics. 2019 Apr;19(8):e1800148.
12. Han H, Cho JW, Lee S et al. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. Nucleic Acids Res. 2018 Jan 4;46(D1):D380-d386.
13. Walker BA, Mavrommatis K, Wardell CP et al. Identification of novel mutational drivers reveals oncogene dependencies in multiple myeloma. Blood. 2018 Aug 9;132(6):587-597.
14. Newman AM, Steen CB, Liu CL et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol. 2019 Jul;37(7):773-782.
15. Patel AP, Tirosh I, Trombetta JJ et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science. 2014 Jun 20;344(6190):1396-401.
16. Zhang H, Meltzer P, Davis S. RCircos: an R package for Circos 2D track plots. BMC Bioinformatics. 2013 Aug 10;14:244.
17. Andreatta M, Carmona SJ. STACAS: Sub-Type Anchor Correction for Alignment in Seurat to integrate single-cell RNA-seq data. Bioinformatics. 2021 May 5;37(6):882-884.
18. Alshetaiwi H, Pervolarakis N, McIntyre LL et al. Defining the emergence of myeloid-derived suppressor cells in breast cancer using single-cell transcriptomics. Sci Immunol. 2020 Feb 21;5(44).
19. Cheng S, Li Z, Gao R et al. A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. Cell. 2021 Feb 4;184(3):792-809.e23.
20. Andreatta M, Corria-Osorio J, Müller S et al. Interpretation of T cell states from single-cell transcriptomics data using reference atlases. Nature communications. 2021 May 20;12(1):2965.
21. Sun Y, Wu L, Zhong Y et al. Single-cell landscape of the ecosystem in early-relapse hepatocellular carcinoma. Cell. 2021 Jan 21;184(2):404-421.e16.
22. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics (Oxford, England). 2010 Jan 1;26(1):139-40.
23. 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 Jul;18(7):1063-1073.
24. Yu Z, Wei X, Liu L et al. Indirubin-3'-monoxime acts as proteasome inhibitor: Therapeutic application in multiple myeloma. EBioMedicine. 2022 Apr;78:103950.

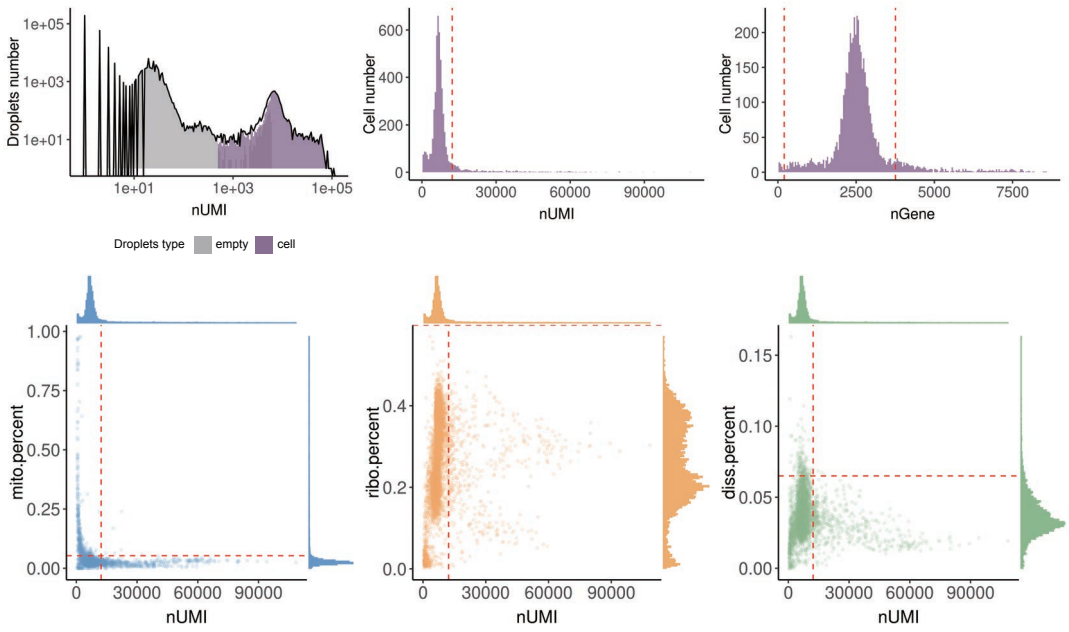
Suppl. Fig. 1

A

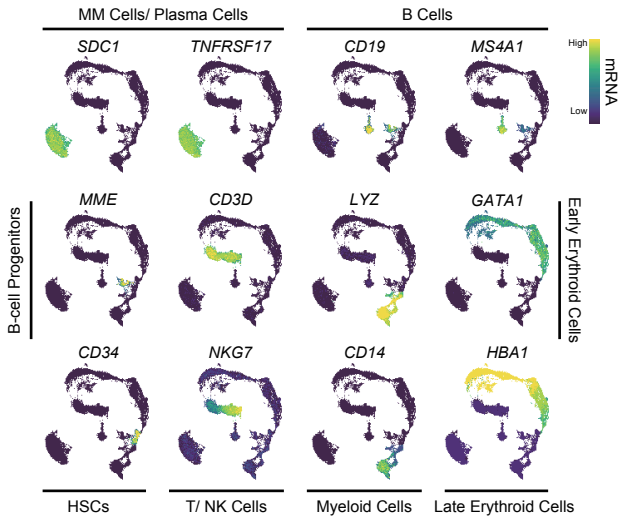
MM sample No.	S8	S12	S15	S24	S1	S2	S3	S4	S5	S16	S25	S27
Myeloma type												
Ig type	IgG	IgA	light chain	IgG	light chain	IgG	IgA	IgA	IgG	IgG	IgA	IgG
light chain type	κ	λ	κ	κ	λ	κ	κ	κ	κ	κ	λ	κ
Internetional Staging System												
ISS III	III	III	III	III	III	I	III	I	II	III	III	III
R-ISS	II	II	III	III	III	I	II	II	II	II	III	III
LDH (High)			■	■								
Cytogenetic aberrancy												
1q+	■	■	■	■	■			■	■		■	
del(17p)			■									
t(4;14)					■			■	■		■	
Treatment class												
VRD regimen	■	■	■	■	■	■	■	■	■	■	■	■
Best response	MR	PR	NR	VGPR	sCR	VGPR	CR	VGPR	PR	PR	CR	CR
Early mortality in 24 months	■	■	■	■								

Red: positive, Blank: negative

B



C



D

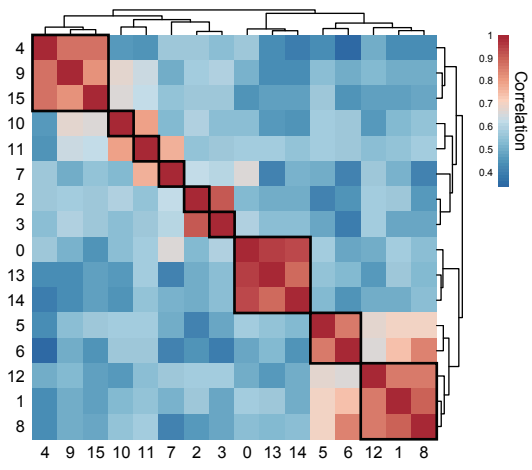


Fig. S1 | The baseline characteristics of MM patients and quality control of scRNA-seq data

(A) Diagram showing the baseline characteristics of multiple myeloma (MM) patients. Red indicates positive and blank indicates negative.

(B) Bar plots showing filtering criteria for sequenced cells, including unique molecular identifier (UMI) count, gene count, mitochondrial gene proportion, ribosomal gene proportion, and percent of dissociation-related genes.

(C) Uniform manifold approximation and projection (UMAP) plots showing the expression of marker genes in the indicated cell types. The labeled bars represent the specific cell types.

(D) Heatmap illustrating the correlation between each cell cluster. The black frame indicates cell clusters of the same cell annotation.

Suppl. Fig. 2

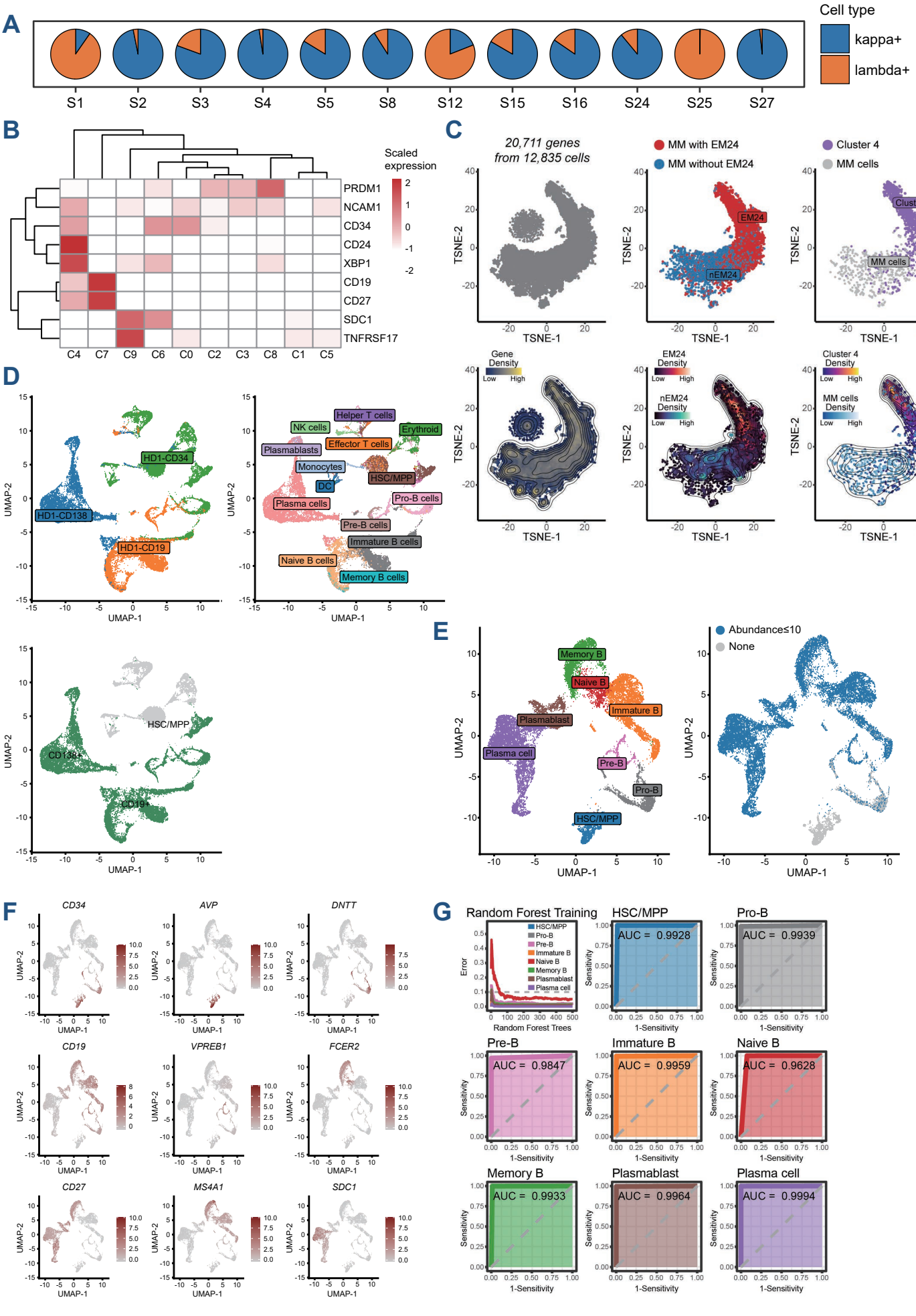

A
Cell type
kappa+
lambda+
S1
S2
S3
S4
S5
S8
S12
S15
S16
S24
S25
S27
B
Scaled expression
PRDM1
NCAM1
CD34
CD24
XBP1
CD19
CD27
SDC1
TNFRSF17
C4
C7
C9
C6
C0
C2
C3
C8
C1
C5
C
20,711 genes from 12,835 cells
MM with EM24
MM without EM24
Cluster 4
MM cells
EM24
nEM24
Gene Density
EM24 Density
nEM24 Density
Cluster 4 Density
MM cells Density
TSNE-1
TSNE-2
D
HD1-CD34
HD1-CD138
HD1-CD19
Helper T cells
NK cells
Erythroid
Effector T cells
Plasmablasts
Monocytes
HSC/MPP
DC
Pro-B cells
Plasma cells
Pre-B cells
Immature B cells
Naive B cells
Memory B cells
CD138+
CD19+
UMAP-1
UMAP-2
E
Memory B
Naive B
Immature B
Plasmablast
Plasma cell
Pre-B
Pro-B
HSC/MPP
Abundance≤10
None
F
CD34
AVP
DNTT
CD19
VPREB1
FCER2
CD27
MS4A1
SDC1
G
Random Forest Training
Random Forest Trees
Error
HSC/MPP
Pro-B
Pre-B
Immature B
Naive B
Memory B
Plasmablast
Plasma cell
AUC = 0.9928
AUC = 0.9939
AUC = 0.9847
AUC = 0.9959
AUC = 0.9628
AUC = 0.9933
AUC = 0.9964
AUC = 0.9994
Sensitivity
1-Sensitivity

Fig. S2 | Plasma-cell analysis of scRNA-seq data

(A) Pie charts showing the proportion of kappa$^+$ and lambda$^+$ plasma cells in each patient.

(B) Heatmap showing the expression of plasma-cell marker gene in each plasma-cell cluster.

(C) T-distributed stochastic neighbor embedding (tSNE) plots (top panel) and density (bottom panel) plots illustrating the topology of sequenced genes from plasma cells. Each point indicates one gene. All genes are shown on the left. The up-regulated genes in patients with early death (EM24) (red) and patients without early death (nEM24) (blue) are shown in the middle. The up-regulated genes in cluster 4 (purple) and other multiple myeloma (MM) cells (gray) are shown on the right.

Uniform manifold approximation and projection (UMAP) plots illustrating cell origin (left), cell annotation (middle), and selected cells for downstream analysis (right) of integrated CD34$^+$, CD19$^+$, and CD138$^+$ cells from integrated outsourced data. HD: healthy individual, HSC: hematopoietic stem cells, MPP: multipotent blood progenitors.

(D) UMAP plots showing the cell annotation (left) and BCR expression pattern (right) of CD34$^+$, CD19$^+$, and CD138$^+$ cells from integrated outsourced data.

(E) UMAP plots showing the expression of marker genes of plasma-cell lineage in integrated CD34$^+$, CD19$^+$, and CD138$^+$ cells.

(F) Line plot and receiver operating characteristic plots showing the performance of plasma-cell lineage classifier by random forest model. AUC: area under the curve.

Suppl. Fig. 3

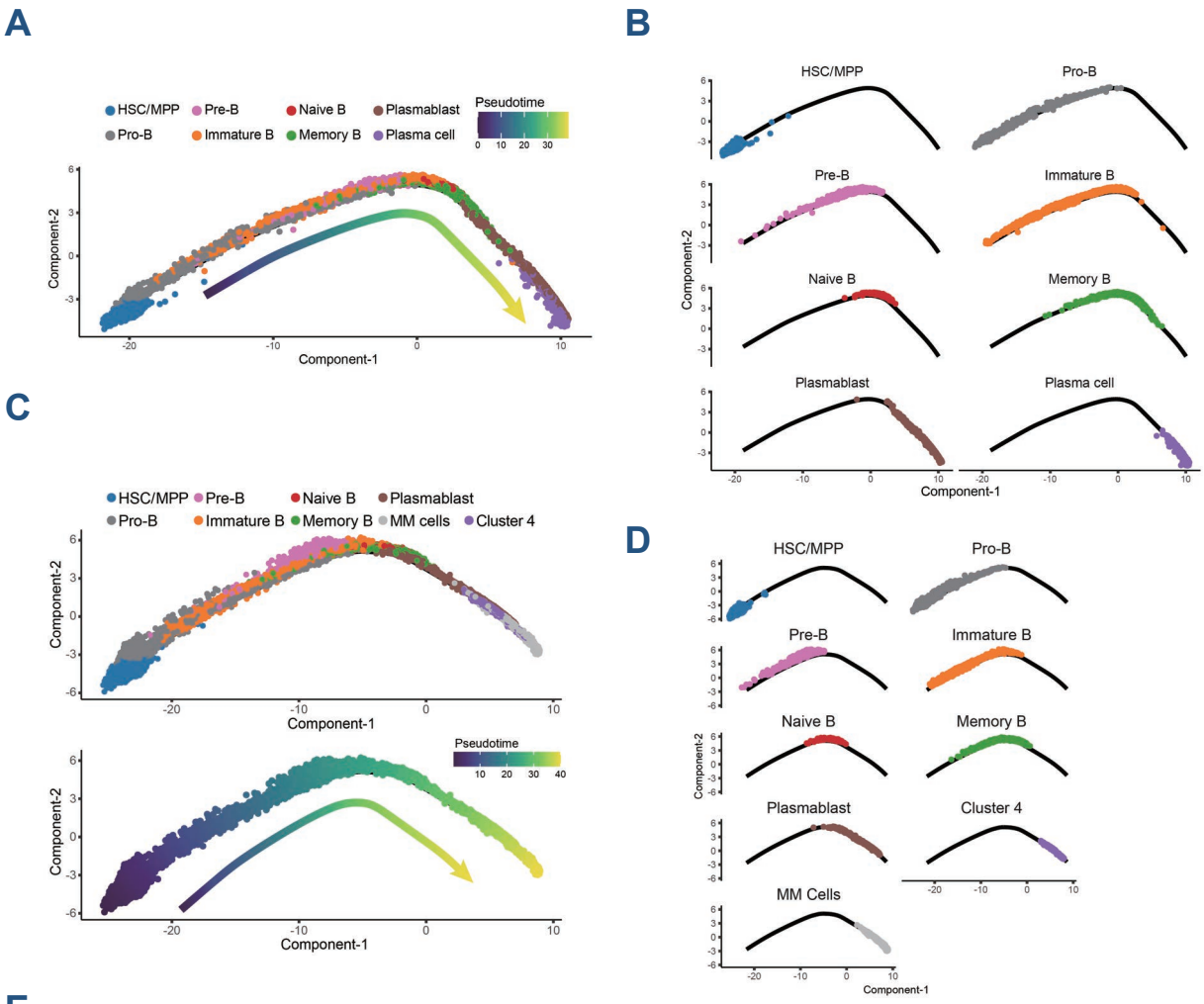

A
HSC/MPP
Pre-B
Naive B
Plasmablast
Pro-B
Immature B
Memory B
Plasma cell
Pseudotime
Component-1
Component-2
B
HSC/MPP
Pro-B
Pre-B
Immature B
Naive B
Memory B
Plasmablast
Plasma cell
C
MM cells
Cluster 4
D
E
REACTOME_SIGNALING_BY_NOTCH
NES: 1.14
FDR: 0.036
REACTOME_SIGNALING_BY_HEDGEHOG
NES: 1.27
FDR: 0.000
GOBP_WNT_SIGNALING_PATHWAY
NES: 1.21
FDR: 0.012
GOBP_STEM_CELL_DIFFERENTIATION
NES: 1.18
FDR: 0.018
Enrichment score (ES)
sC4
others

Fig. S3 | Plasma-cell differentiation trajectory analysis

(A) Pseudotime-ordered trajectory of plasma-cell lineage using integrated $CD34^+$, $CD19^+$, and $CD138^+$ cells. HSC: hematopoietic stem cells, MPP: multipotent blood progenitors.

(B) The distribution of plasma-cell lineage during the trajectory using integrated $CD34^+$, $CD19^+$, and $CD138^+$ cells. Cells are labeled by colors.

(C) Pseudotime-ordered trajectory of plasma-cell lineage involved cluster 4 and other multiple myeloma (MM) cells. Plasma-cell lineage (top) and pseudotime order (bottom) are labeled by colors.

(D) The distribution of MM-cell clusters during the plasma-cell lineage differentiation trajectory. Cells are labeled by colors.

(E) Enrichment plots showing the enriched stemness-related signaling pathways in sub-C4.

Suppl. Fig. 4

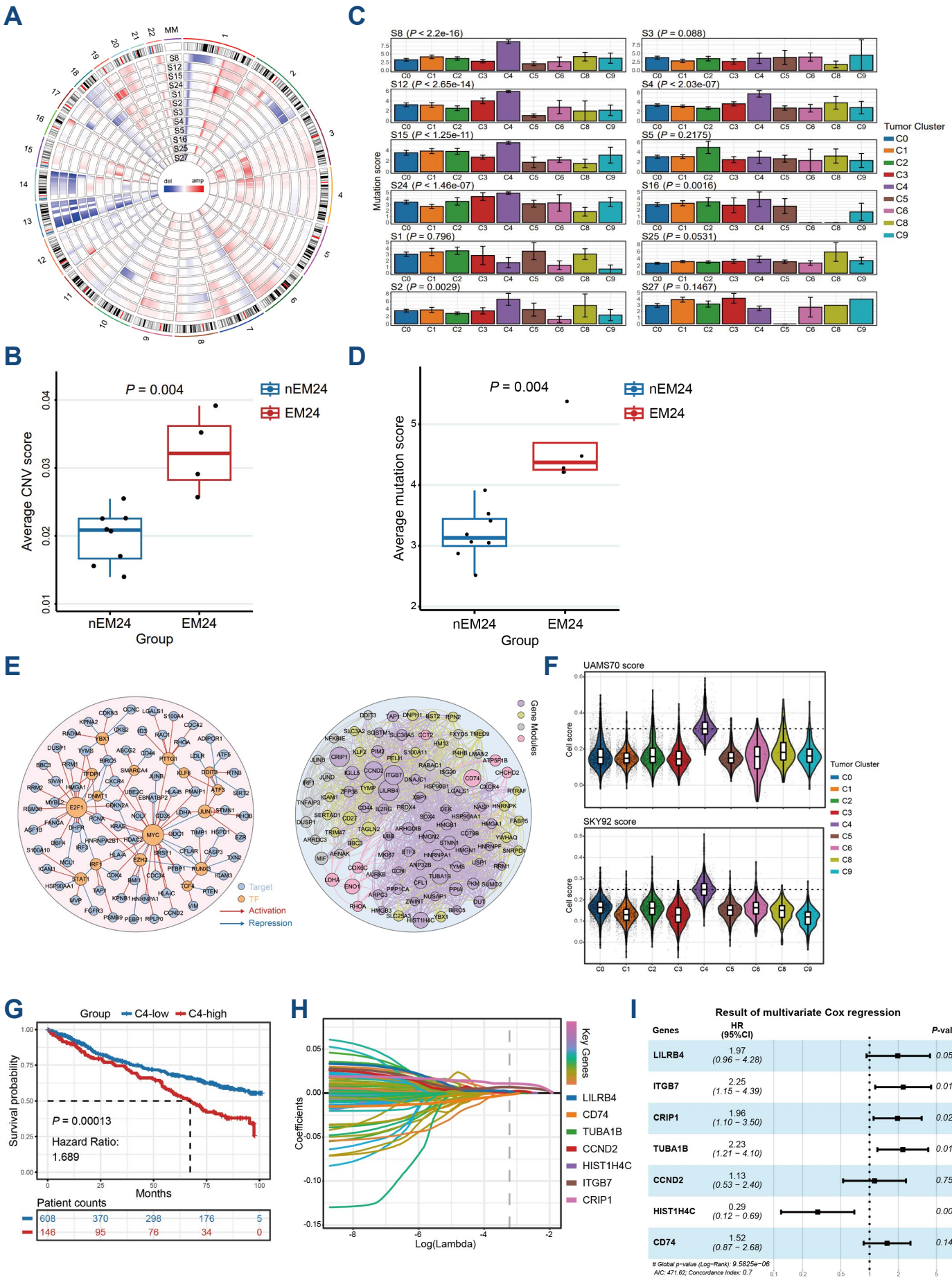

A
C
B
D
E
F
G
H
I
Tumor Cluster
C0
C1
C2
C3
C4
C5
C6
C8
C9
Mutation score
Average CNV score
Average mutation score
P = 0.004
nEM24
EM24
Group
Target
TF
Activation
Repression
Gene Modules
UAMS70 score
SKY92 score
Cell score
Group
C4-low
C4-high
Survival probability
P = 0.00013
Hazard Ratio:
1.689
Months
Patient counts
608
370
298
176
5
146
95
76
34
0
Coefficients
Log(Lambda)
Key Genes
LILRB4
CD74
TUBA1B
CCND2
HIST1H4C
ITGB7
CRIP1
Result of multivariate Cox regression
Genes
HR (95%CI)
P-value
LILRB4 1.97 (0.96 – 4.28) 0.057
ITGB7 2.25 (1.15 – 4.39) 0.018 *
CRIP1 1.96 (1.10 – 3.50) 0.022 *
TUBA1B 2.23 (1.21 – 4.10) 0.01 *
CCND2 1.13 (0.53 – 2.40) 0.755
HIST1H4C 0.29 (0.12 – 0.69) 0.006 **
CD74 1.52 (0.87 – 2.68) 0.142
Global p-value (Log-Rank): 9.5825e-06
AIC: 471.62; Concordance Index: 0.7

Fig. S4 | Genomic features of sub-cluster 4

(A) Circular genomic map illustrating the copy number variations (CNV) in each multiple myeloma (MM) patient. Each lane indicates one patient and the outmost lane indicates chromosome structure. Red represents amplification and blue represents deletion.

(B) Box plot showing the average CNV score between patients with early death (EM24) and patients without early death (nEM24).

(C) Bar plots showing the mutation score of each tumor cluster in each patient.

(D) Box plot showing the mutation score between EM24 patients and nEM24 patients.

(E) (Left) Regulatory network of major transcription factors (TF) and target genes up-regulated in cluster 4 compared with other MM cells. The point size indicates the edge numbers of TF. (Right) Gene co-expression network displaying the top gene modules specific to cluster 4. The point size indicates the gene fold change in cluster 4 versus other MM cells. The lines connect significantly correlated genes.

(F) Violin plots showing the cell score of UAMS70 and SKY92 in MM cell clusters.

(G) Kaplan-Meier curve showing the overall survival of 754 MM patients with a high or low proportion of sub-C4 in the MMRF-CoMMpass dataset. Log-rank test was applied in the comparison between groups.

(H) Line chart showing the least absolute shrinkage and selection operator (lasso) coefficient profiles of the 170 risk factors against log(Lambda).

(I) Forest plot showing the hazard ratio of 7 genes by multivariate cox regression in MMRF-CoMMpass cohort.

Suppl. Fig. 5

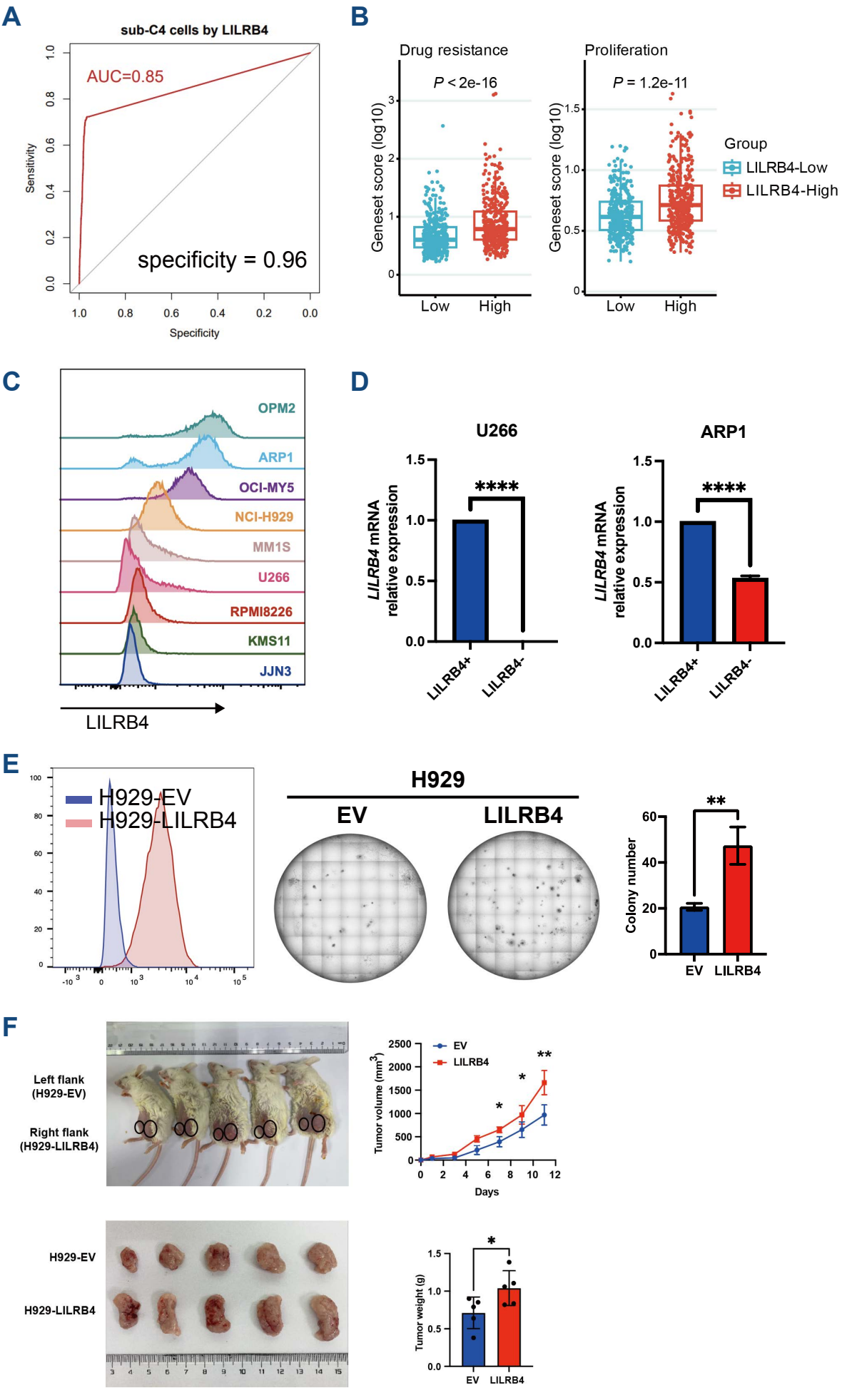

A
sub-C4 cells by LILRB4
AUC=0.85
specificity = 0.96
Sensitivity
Specificity
B
Drug resistance
P < 2e-16
Proliferation
P = 1.2e-11
Geneset score (log10)
Low
High
Group
LILRB4-Low
LILRB4-High
C
OPM2
ARP1
OCI-MY5
NCI-H929
MM1S
U266
RPMI8226
KMS11
JJN3
LILRB4
D
U266
ARP1
LILRB4 mRNA relative expression
****
LILRB4+
LILRB4-
E
H929-EV
H929-LILRB4
H929
EV
LILRB4
Colony number
**
F
Left flank (H929-EV)
Right flank (H929-LILRB4)
Tumor volume (mm3)
Days
H929-EV
H929-LILRB4
Tumor weight (g)

Fig. S5 | LILRB4 is a biomarker of aggressive MM

(A) Receiver operating characteristic (ROC) curve showing the specificity of LILRB4 expression in sub-C4. AUC: area under the curve.

(B) The score of drug-resistance and proliferation geneset between LILRB4-low and LILRB4-high patients in MMRF-CoMMpass cohort.

(C) Flow cytometry detection of LILRB4 in multiple multiple myeloma (MM) cell lines.

(D) LILRB4 mRNA expression of $LILRB4^{+}$ and $LILRB4^{-}$ cells. Unpaired *t* test, two-tailed, ****P<0.0001

(E) (Left) Flow cytometry detection of LILRB4 in H929-EV and H929-LILRB4 cells. (Right) Colony formation assay of H929-LILRB4 cells relative to H929-EV cells. Unpaired *t* test, two-tailed, **P<0.01.

(F) The H929-EV and H929-LILRB4 cells were injected subcutaneously into the left and right flank of the same mouse. Bar plots showing the statistical result of tumor weight between H929-EV and H929-LILRB4 group with unpaired *t* test, two-tailed, *P < 0.05, **P<0.01.

Suppl. Fig. 6

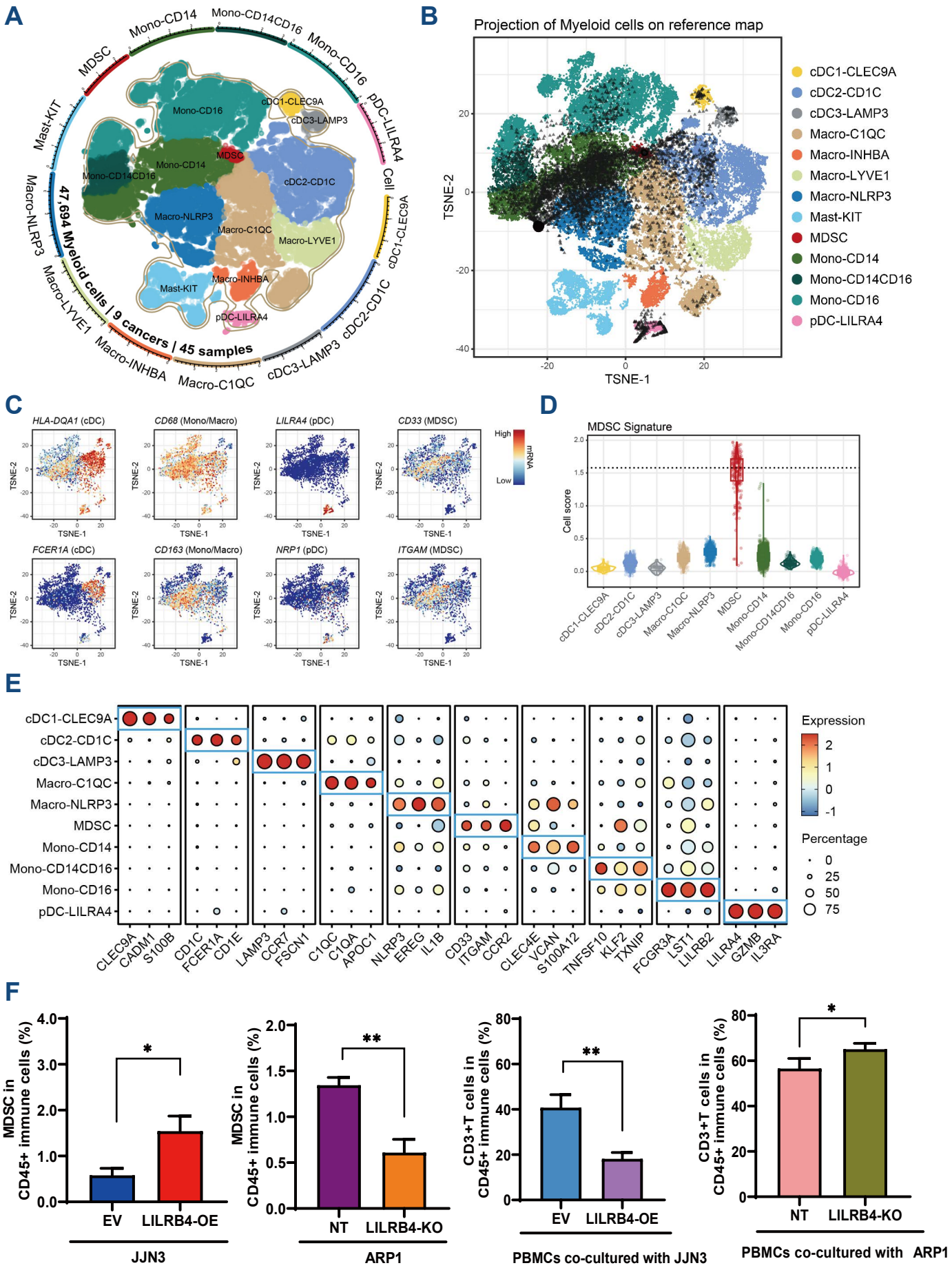

A
Mono-CD14
Mono-CD14CD16
Mono-CD16
pDC-LILRA4
cDC1-CLEC9A
cDC2-CD1C
cDC3-LAMP3
Macro-C1QC
Macro-INHBA
Macro-LYVE1
Macro-NLRP3
Mast-KIT
MDSC
47,694 Myeloid cells | 9 cancers | 45 samples
B
Projection of Myeloid cells on reference map
TSNE-1
TSNE-2
C
HLA-DQA1 (cDC)
CD68 (Mono/Macro)
LILRA4 (pDC)
CD33 (MDSC)
FCER1A (cDC)
CD163 (Mono/Macro)
NRP1 (pDC)
ITGAM (MDSC)
High
Low
mRNA
D
MDSC Signature
Cell score
E
Expression
Percentage
CLEC9A CADM1 S100B CD1C FCER1A CD1E LAMP3 CCR7 FSCN1 C1QC C1QA APOC1 NLRP3 EREG IL1B CD33 ITGAM CCR2 CLEC4E VCAN S100A12 TNFSF10 KLF2 TXNIP FCGR3A LST1 LILRB2 LILRA4 GZMB IL3RA
F
MDSC in CD45+ immune cells (%)
EV
LILRB4-OE
JJN3
NT
LILRB4-KO
ARP1
CD3+T cells in CD45+ immune cells (%)
PBMCs co-cultured with JJN3
PBMCs co-cultured with ARP1

Fig. S6 | Myeloid cell annotation and signature gene expression

(A) Points plot showing integrated myeloid cell map for reference, including 47,694 cells from 45 samples of 9 cancers.

(B) T-distributed stochastic neighbor embedding (tSNE) plot showing the distribution of annotated myeloid cells. Cell annotations are labeled by colors.

(C) TSNE plots showing the marker gene expression for the major lineages of myeloid cells.

(D) Violin plot showing the myeloid-derived suppressive cell (MDSC) signature scores in myeloid cell types.

(E) Bubble heatmap showing expression levels of selected signature genes in myeloid cells. Dot size indicates fraction of expressing cells, colored based on normalized expression levels.

(F) Bar plots showing the proportion of MDSCs or $CD3^+$ T cells in peripheral blood mononuclear cells (PBMCs) after co-culturing with multiple myeloma (MM) cells with different LILRB4 level respectively. The statistical results were tested with unpaired *t* test, two-tailed, $*P < 0.05$, $**P<0.01$.

Suppl. Fig. 7

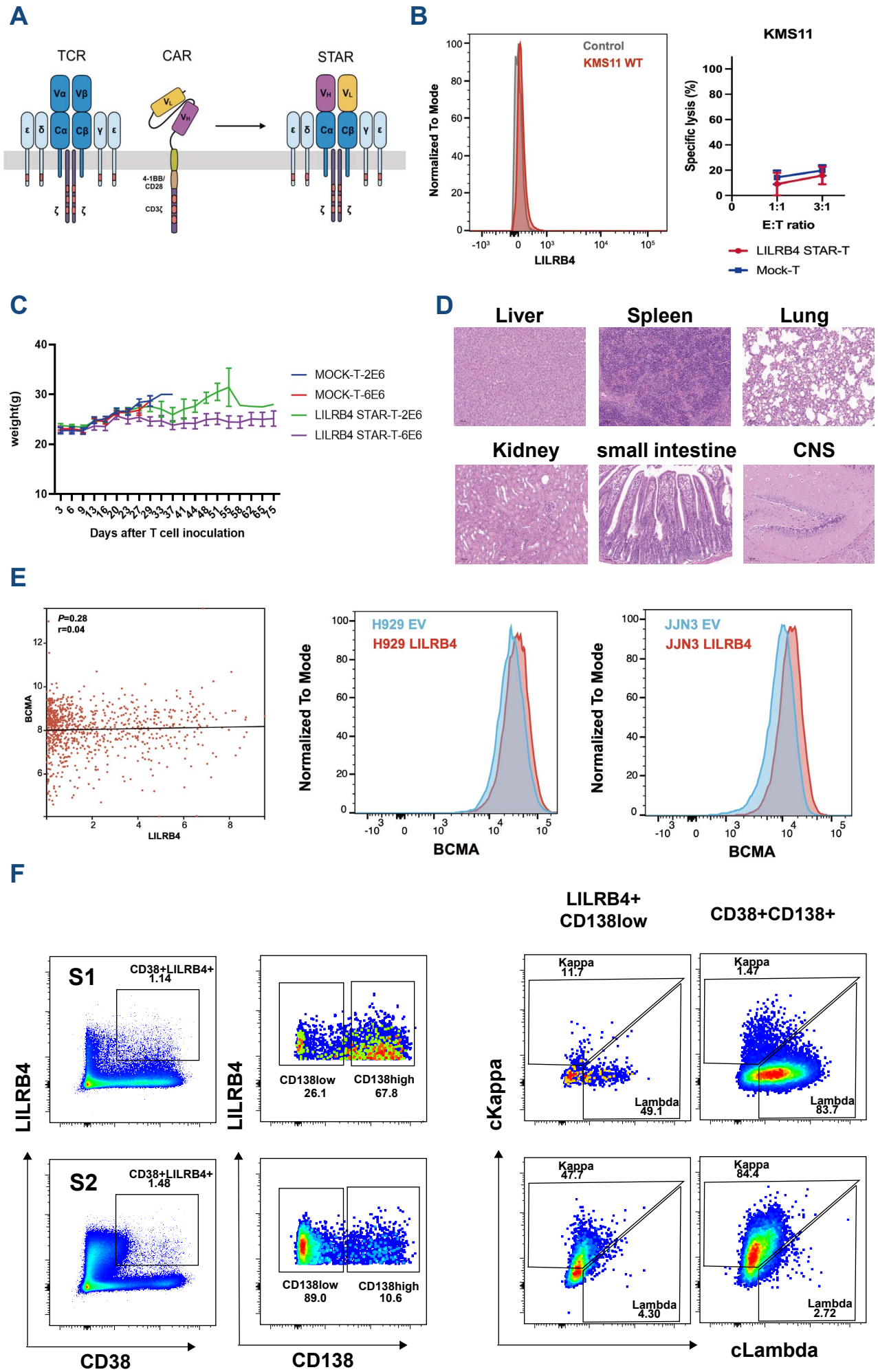

A
TCR
CAR
STAR
Vα
Vβ
Cα
Cβ
ε
δ
γ
ζ
VL
VH
4-1BB/
CD28
CD3ζ
B
Control
KMS11 WT
Normalized To Mode
LILRB4
KMS11
Specific lysis (%)
E:T ratio
1:1
3:1
LILRB4 STAR-T
Mock-T
C
weight(g)
Days after T cell inoculation
MOCK-T-2E6
MOCK-T-6E6
LILRB4 STAR-T-2E6
LILRB4 STAR-T-6E6
D
Liver
Spleen
Lung
Kidney
small intestine
CNS
E
P=0.28
r=0.04
BCMA
LILRB4
H929 EV
H929 LILRB4
JJN3 EV
JJN3 LILRB4
Normalized To Mode
BCMA
F
LILRB4+
CD138low
CD38+CD138+
S1
CD38+LILRB4+
1.14
CD138low
26.1
CD138high
67.8
Kappa
11.7
Lambda
49.1
Kappa
1.47
Lambda
83.7
S2
CD38+LILRB4+
1.48
CD138low
89.0
CD138high
10.6
Kappa
47.7
Lambda
4.30
Kappa
84.4
Lambda
2.72
CD38
CD138
cKappa
cLambda

Fig. S7 | LILRB4 STAR-T cell pre-clinical study in MM

(A) Schematic construct of LILRB4 synthetic T cell receptor and antigen receptor (STAR)-T cells.

(B) LILRB4 STAR-T cell cytotoxicity assay towards LILRB4$^-$ KMS11 multiple myeloma (MM) cells.

(C) Line chart illustrating the subcutaneous tumor weight in MM xenograft mice after treating with mock-T cells or LILRB4 STAR-T cells.

(D) Hematoxylin-eosin staining of multiple organs from MM xenograft mice after treating with LILRB4 STAR-T cells.

(E) (Left) Correlation analysis of *BCMA* and *LILRB4* mRNA expression in GSE2658 dataset. (Right) Flow cytometry detection of BCMA in LILRB4-overexpressed MM cells.

(F) Exemplary plot for light chain restriction analysis of CD38$^+$LILRB4$^+$CD138^{low} cell cluster and CD38$^+$CD138$^+$ cell cluster from the same patient by flow cytometry detection.