Supplementary Methods

1. Sample Preparation

Fresh bone marrow (BM) aspirate samples were cryopreserved in 90% fetal bovine serum and subsequently sorted into CD138+ and CD138- fractions. Samples were stored in liquid nitrogen until ready for use.

2.Tissue Microarray Analysis by Multiplexed Immunofluorescence

First, cell segmentation was carried out on the DAPI stained nuclei using the software's nuclei detection application. A region of interest was selected for each core for the evaluation of each marker which included positive cells, negative cells and the background area. This was done to train the software to differentiate between marker-specific positive and negative cells. Then, individual deep learning APPs (Application Protocol Packages) were created for all markers and the data was analysed via batch analysis. The data generated included the total number of nuclei, mean intensity for each marker and object information for all cells in each core. The object information labels each cell as positive or negative for a particular marker. When this data was compiled and exported, we were able to identify specific cell populations.

3.Hyperplex Immunohistochemistry for Spatial Phenotyping

The images were analysed using HALO v3.6 (Indica Labs, Albuquerque, NM, USA) software. Binary classification of each protein biomarker was conducted based on visually-guided thresholding of intensity values. Then, proximity analysis was carried out to tabulate the distance (within 500μm)(1) between major immune cell types and CD38+CD138+ plasma cells was carried out. Using RStudio (Ver. 2023.6.1.524), the distribution of the distances between each major immune cell type and CD38+CD138+ plasma cells was visualized and the average distance was tabulated. Using RStudio (Ver. 2023.6.1.524),the SPIAT package (https://trigosteam.github.io/SPIAT/) was used for spatial analysis. Clusters were defined based on manual phenotyping of the protein biomarkers, and mean pairwise distance was tabulated between the clusters, taking into account all cells and distances. Lastly, the Wilcoxon rank sum test was used to test for any statistical significance between the mean pairwise distances of immune cell clusters and plasma cell cluster.

4.Pre-processing of Single Cell Sequencing Data

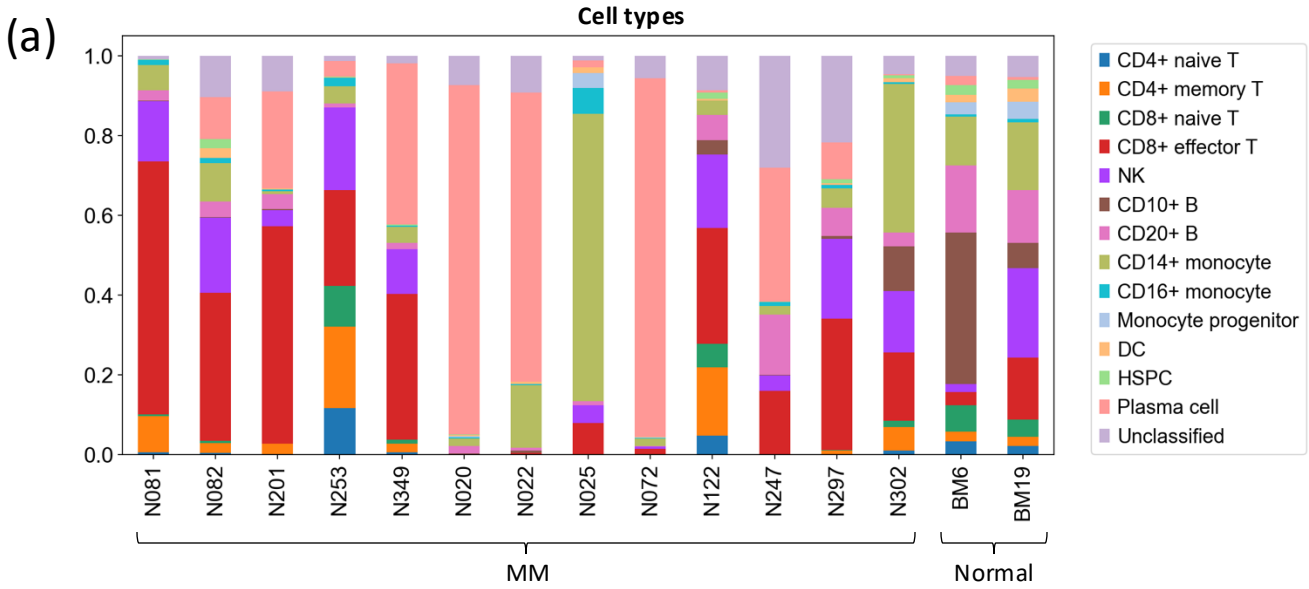
We pre-processed raw single-cell sequencing data using the 10X Cell Ranger software, and generated single-cell gene count matrices. We used the SoupX R package (2) to estimate and remove cell-free mRNA contamination. Thereafter, we processed and analyzed the data in Python, using toolkits including Scanpy and MapBatch. We removed putative doublets and empty cells as cells with total UMI counts greater than or less than 3 standard deviations from the mean counts within each batch, and removed potential dead cells as cells with percentage of mitochondrial RNA greater than 10%. Data was normalized by library size followed by log-scaling (log normalization).

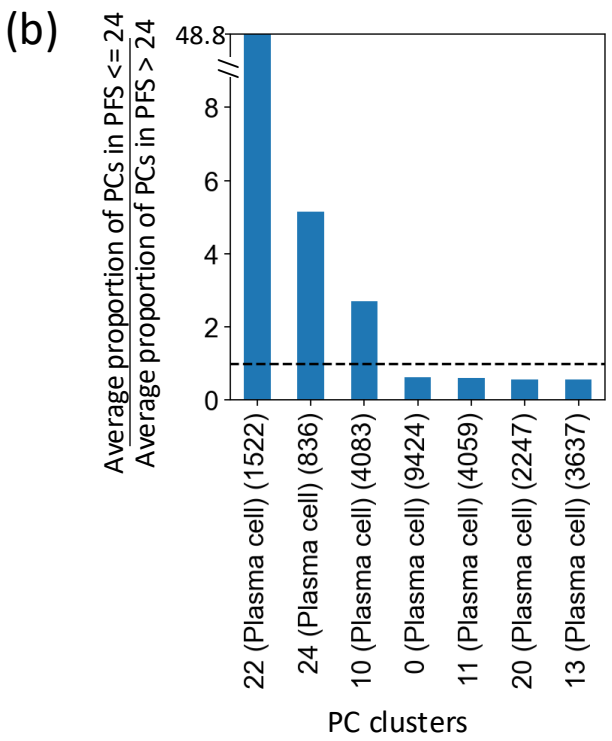
The public dataset subset of the validation cohort was downloaded from the single-cell MM atlas (3) but only the subset published by Maura et al. 2023 (4) was used as the others lacked progression data. The data was downloaded in annotated gene-counts format, and the data was log-normalized. This was combined with our generated subset, including only the genes both subsets had in common (26284 genes).Single-cell protein expression data was normalized using CLR normalization, and for each protein, a histogram of expression counts was plotted to determine a cut-off threshold for expression.

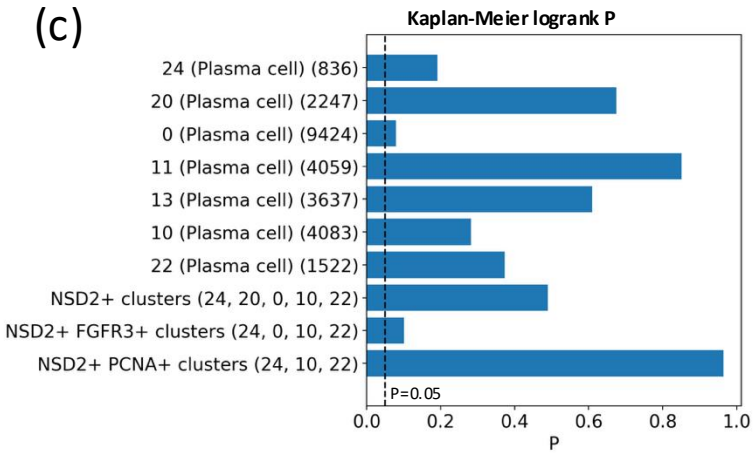
References

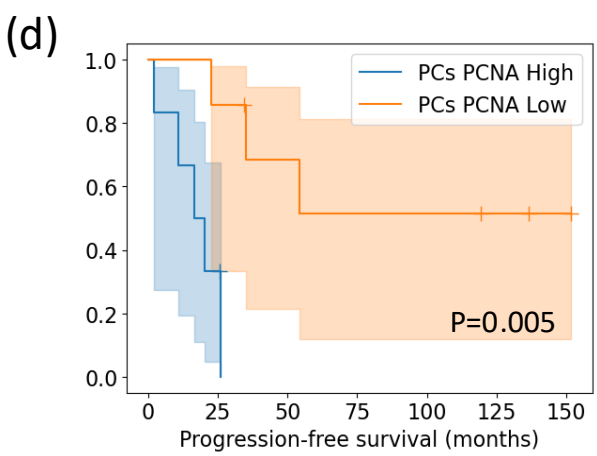
1. Feng Y, Yang T, Zhu J, Li M, Doyle M, Ozcoban V, et al. Spatial analysis with SPIAT and spaSim to characterize and simulate tissue microenvironments. Nature Communications. 2023;14(1):2697.
2. Young MD, Behjati S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. GigaScience. 2020;9(12).
3. Foster KA, Rees E, Ainley L, Boyle EM, Lee L, Ward G, et al. Tumour-intrinsic features shape T-cell differentiation through myeloma disease evolution. medRxiv. 2024:2024.06.22.24309250.
4. Maura F, Boyle EM, Coffey D, Maclachlan K, Gagler D, Diamond B, et al. Genomic and immune signatures predict clinical outcome in newly diagnosed multiple myeloma treated with immunotherapy regimens. Nat Cancer. 2023;4(12):1660-74.

Supplementary Figure 1 (accompanies Figure 1). (a) Breakdown of cell types in each sample. **(b)** Comparative prevalence of PC clusters among poorer-PFS versus better-PFS patients. Ratio>1 indicates the cluster is more prevalent in PFS≤24 samples; ratio<1 indicates the cluster is more prevalent in PFS>24 samples. **(c)** Logrank test P values for associations between PC clusters and *NSD2*⁺/*FGFR3*⁺/*PCNA*⁺ cluster combinations, and PFS. **(d)** KM survival analysis of the PC proliferative index (proportion of *PCNA*⁺ cells among PCs) in PCNA High and Low groups.



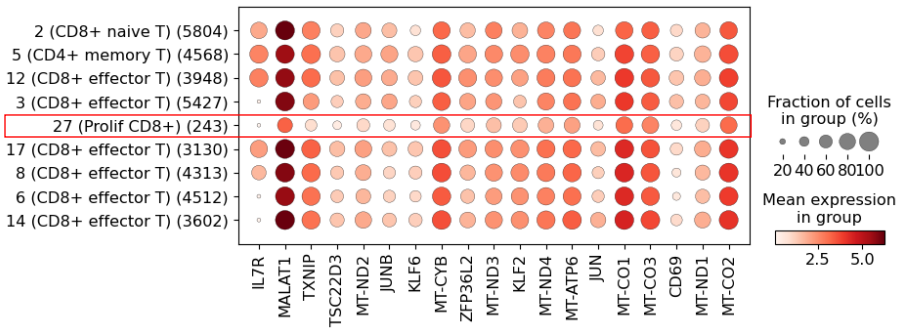




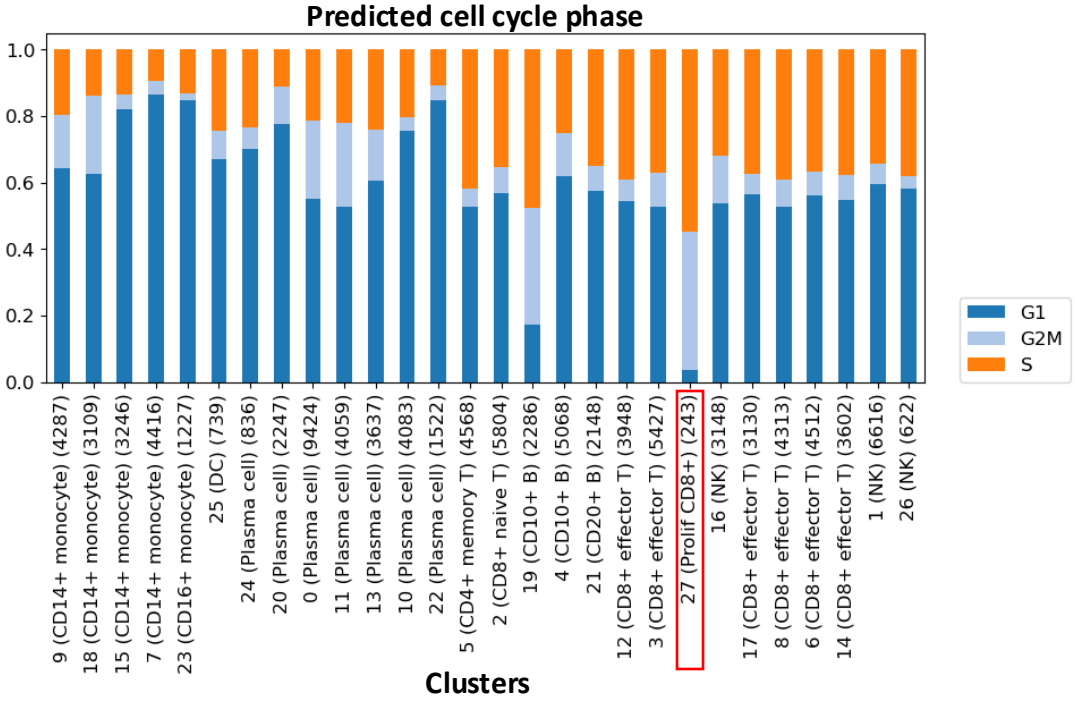


Supplementary Figure 2 (accompanies Figure 2). (a) Downregulated genes in Cluster 27 cells compared to other T cells. **(b)** Cell cycle phase prediction of TME clusters, demonstrating the proportion of cells in each phase of the cell cycle by cluster. **(c)** Cell identification prediction of cell clusters using DISCO CellID.

(a)



(b)



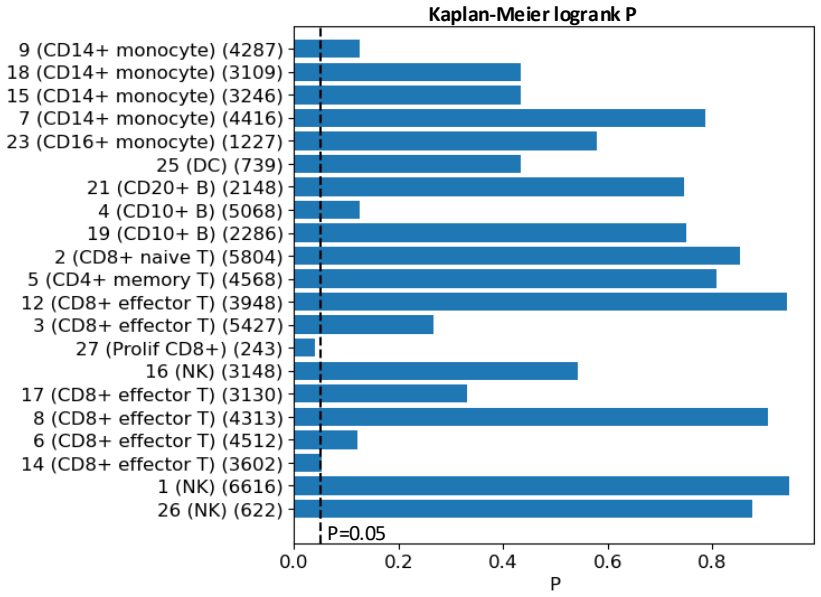
(c)

DISCO CellID

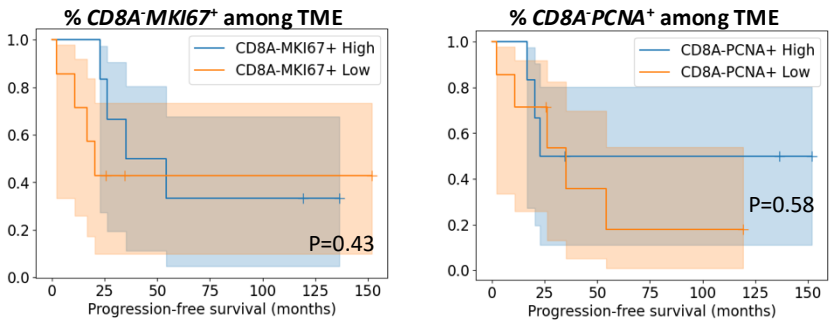
Cluster	Primary Prediction	Primary Source Atlas	Primary Score	Secondary Prediction	Secondary Source Atlas	Secondary Score
0 (Plasma cell) (9424)	Plasma cell--tonsil	Plasma cell--tonsil	0.56	Class-switched memory B cell--tonsil	Class-switched memory B cell--tonsil	0.408
1 (NK) (6616)	CD16 NK cell--blood	CD16 NK cell--blood	0.914	CD16 NK cell--sarcoidosis_blood	CD16 NK cell--sarcoidosis_blood	0.902
2 (CD8+ naive T) (5804)	Naive CD4 T cell--blood	Naive CD4 T cell--blood	0.899	Naive CD4 T cell--sarcoidosis_blood	Naive CD4 T cell--sarcoidosis_blood	0.895
3 (CD8+ effector T) (5427)	GZMK CD8 T cell--bone_marrow	GZMK CD8 T cell--bone_marrow	0.902	GZMK CD8 T cell--blood	GZMK CD8 T cell--blood	0.864
4 (CD10+ B) (5068)	Naive B cell--bone_marrow	Naive B cell--bone_marrow	0.785	Memory B cell--bone_marrow	Memory B cell--bone_marrow	0.765
5 (CD4+ memory T) (4568)	Memory CD4 T cell--sarcoidosis_blood	Memory CD4 T cell--sarcoidosis_blood	0.915	Memory CD4 T cell--HNSCC_blood	Memory CD4 T cell--HNSCC_blood	0.915
6 (CD8+ effector T) (4512)	GZMB CD8 T cell--blood	GZMB CD8 T cell--blood	0.906	GZMB CD8 T cell--HNSCC_blood	GZMB CD8 T cell--HNSCC_blood	0.904
7 (CD14+ monocyte) (4416)	CD14 monocyte--HNSCC_blood	CD14 monocyte--HNSCC_blood	0.931	CD14 monocyte--sarcoidosis_blood	CD14 monocyte--sarcoidosis_blood	0.915
8 (CD8+ effector T) (4313)	GZMB CD8 T cell--blood	GZMB CD8 T cell--blood	0.908	GZMB CD8 T cell--sarcoidosis_blood	GZMB CD8 T cell--sarcoidosis_blood	0.899
9 (CD14+ monocyte) (4287)]	CD14 monocyte--HNSCC_blood	CD14 monocyte--HNSCC_blood	0.877	CD14 monocyte--sarcoidosis_blood	CD14 monocyte--sarcoidosis_blood	0.851
10 (Plasma cell) (4083)	Plasma cell--Crohns_disease_ileum	Plasma cell--Crohns_disease_ileum	0.815	Plasma cell--tonsil	Plasma cell--tonsil	0.693
11 (Plasma cell) (4059)	Plasma cell--tonsil	Plasma cell--tonsil	0.623	Memory CD4 T cell--tonsil	Memory CD4 T cell--tonsil	0.586
12 (CD8+ effector T) (3948)	Memory CD4 T cell--HNSCC_blood	Memory CD4 T cell--HNSCC_blood	0.888	GZMK CD8 T cell--blood	GZMK CD8 T cell--blood	0.885
13 (Plasma cell) (3637)	Plasma cell--bone_marrow	Plasma cell--bone_marrow	0.808	Plasma cell--thymus	Plasma cell--thymus	0.695
14 (CD8+ effector T) (3602)	GZMB CD8 T cell--blood	GZMB CD8 T cell--blood	0.908	GZMB CD8 T cell--HNSCC_blood	GZMB CD8 T cell--HNSCC_blood	0.902
15 (CD14+ monocyte) (3246)	CD14 monocyte--HNSCC_blood	CD14 monocyte--HNSCC_blood	0.9	CD14 monocyte--sarcoidosis_blood	CD14 monocyte--sarcoidosis_blood	0.882
16 (NK) (3148)	GZMB CD8 T cell--blood	GZMB CD8 T cell--blood	0.825	MAIT cell--blood	MAIT cell--blood	0.822
17 (CD8+ effector T) (3130)	GZMB CD8 T cell--HNSCC_blood	GZMB CD8 T cell--HNSCC_blood	0.916	GZMB CD8 T cell--blood	GZMB CD8 T cell--blood	0.907
18 (CD14+ monocyte) (3109)	CD14 monocyte--HNSCC_blood	CD14 monocyte--HNSCC_blood	0.91	CD14 monocyte--sarcoidosis_blood	CD14 monocyte--sarcoidosis_blood	0.886
19 (CD10+ B) (2286)	Multipotent progenitor--bone_marrow	Multipotent progenitor--bone_marrow	0.833	Lymphoid pre-pDC--bone_marrow	Lymphoid pre-pDC--bone_marrow	0.825
20 (Plasma cell) (2247)	Plasma cell--Crohns_disease_ileum	Plasma cell--Crohns_disease_ileum	0.709	Plasma cell--tonsil	Plasma cell--tonsil	0.617
21 (CD20+ B) (2148)	Memory B cell--bone_marrow	Memory B cell--bone_marrow	0.907	Class-switched memory B cell--tonsil	Class-switched memory B cell--tonsil	0.851
22 (Plasma cell) (1522)	Plasma cell--HNSCC_blood	Plasma cell--HNSCC_blood	0.758	Plasma cell--Crohns_disease_ileum	Plasma cell--Crohns_disease_ileum	0.637
23 (CD16+ monocyte) (1227)	CD16 monocyte--HNSCC_blood	CD16 monocyte--HNSCC_blood	0.934	CD16 monocyte--sarcoidosis_blood	CD16 monocyte--sarcoidosis_blood	0.927
24 (Plasma cell) (836)	CD4 T cell--gingiva	CD4 T cell--gingiva	0.759	Memory CD8 T cell--tonsil	Memory CD8 T cell--tonsil	0.755
25 (DC) (739)	pDC--HNSCC_blood	pDC--HNSCC_blood	0.863	Dendritic cell--blood	Dendritic cell--blood	0.795
26 (NK) (622)	CD56 NK cell--blood	CD56 NK cell--blood	0.888	CD56 NK cell--HNSCC_blood	CD56 NK cell--HNSCC_blood	0.883
27 (Prolif CD8+) (243)	Cycling T/NK cell--HNSCC_blood	Cycling T/NK cell--HNSCC_blood	0.913	Cycling T/NK cell--sarcoidosis_blood	Cycling T/NK cell--sarcoidosis_blood	0.901

Supplementary Figure 2 (continued). (d) Logrank test P values for association between TME cell clusters and PFS. **(e)** Negative control KM survival analyses of *CD8A*$^{-}$*MKI67*$^{+}$ and *CD8A*$^{-}$*PCNA*$^{+}$ High and Low groups. **(f)** Evaluating the correlation between Cluster 27 proportion in the TME and *PCNA*$^{+}$ proportion in PCs.

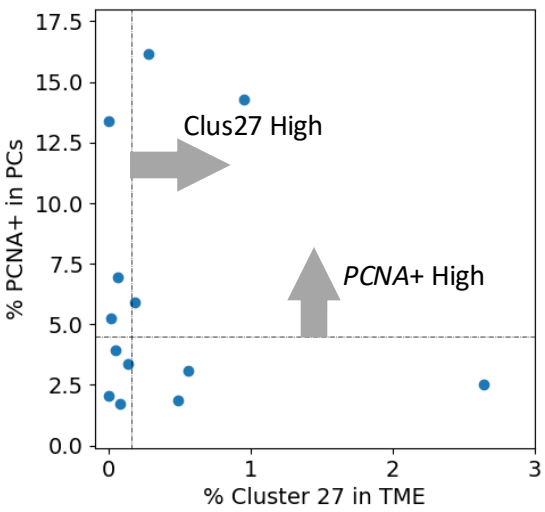
(d)



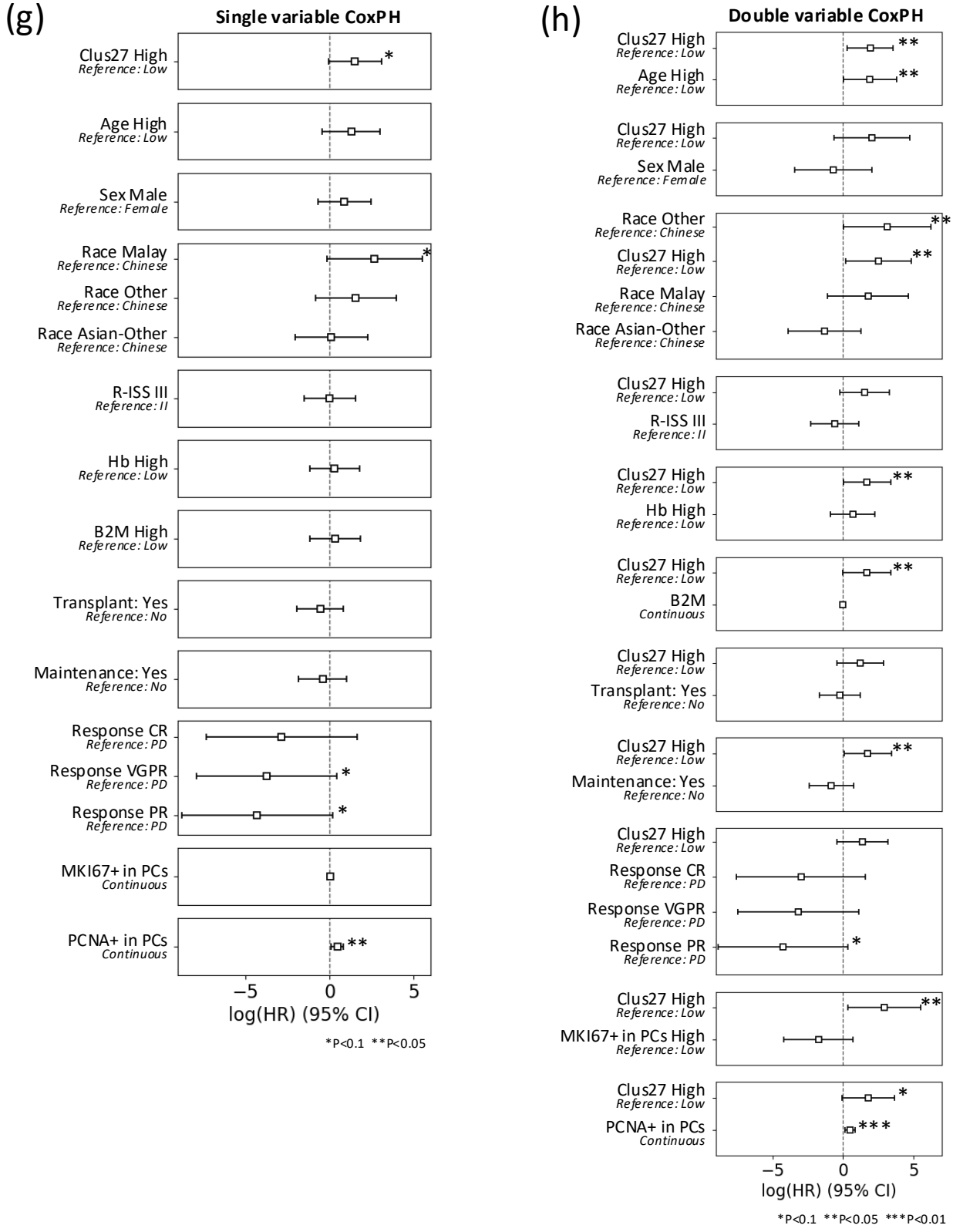
(e)



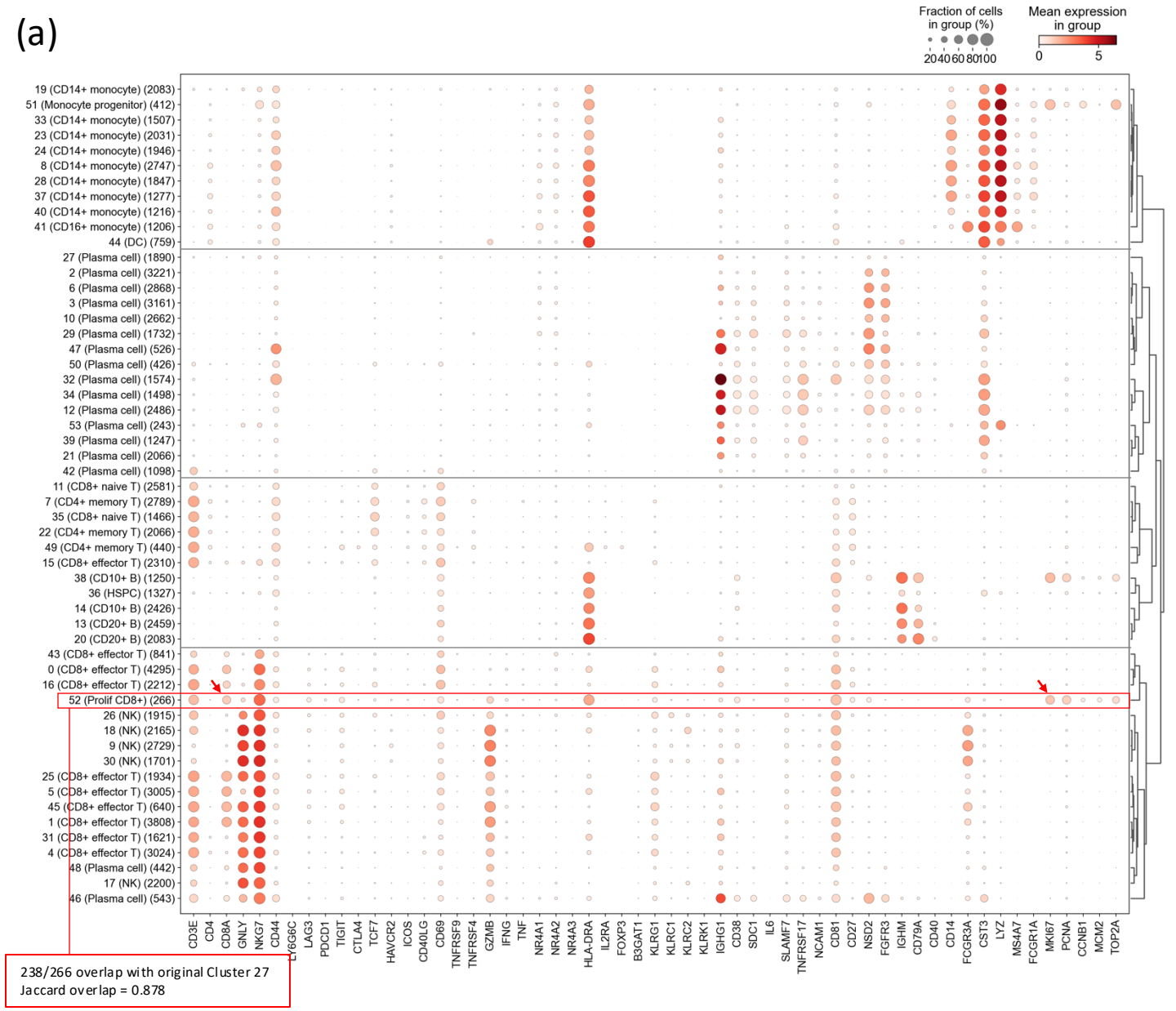
(f)



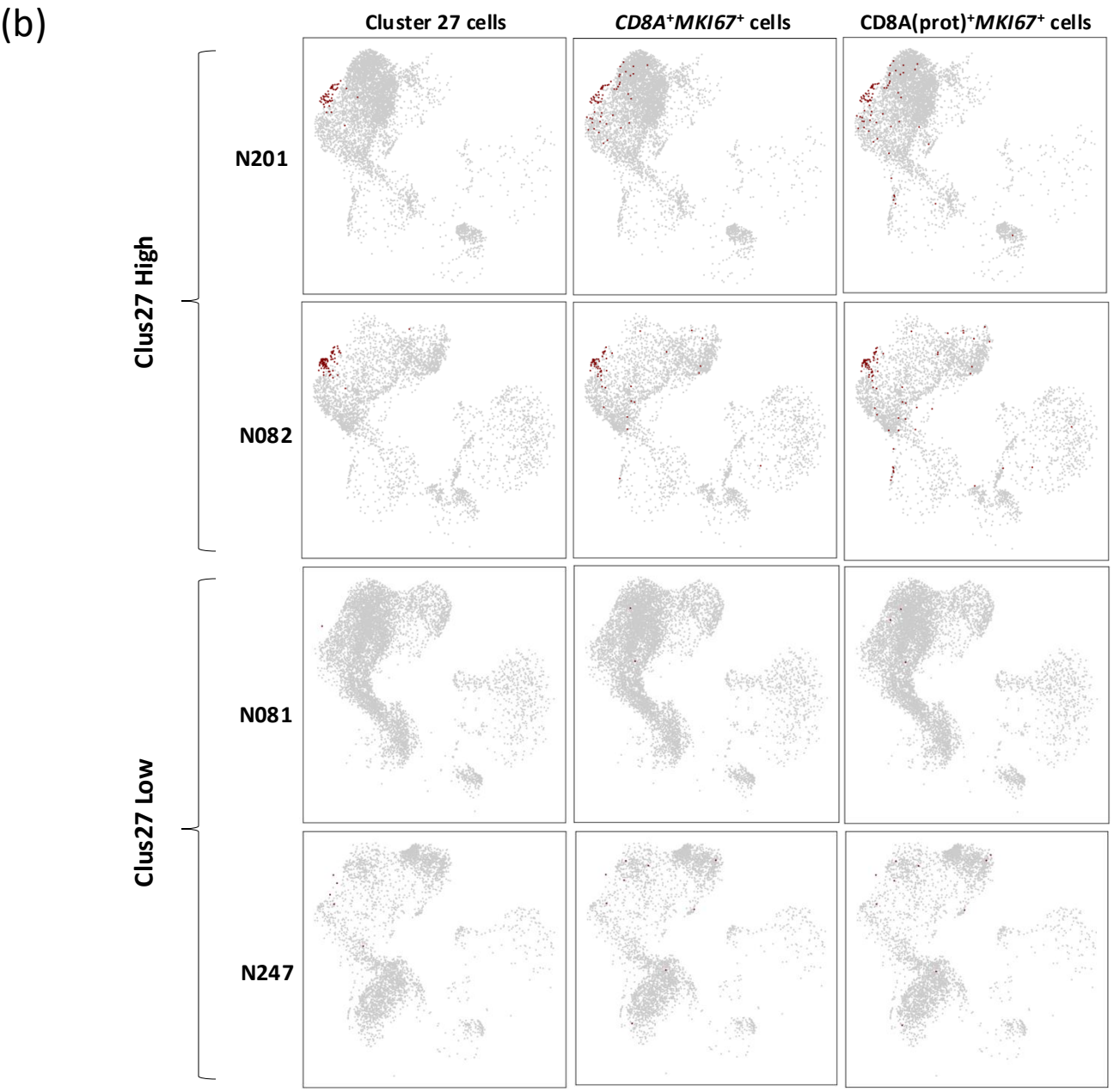
Supplementary Figure 2 (continued). (g) Single-variable CoxPH survival analysis. **(h)** Double-variable CoxPH survival analyses (Clus27 status with each other covariate). Covariates with P<0.1 in single or double-variable analysis were selected for multivariable analysis.



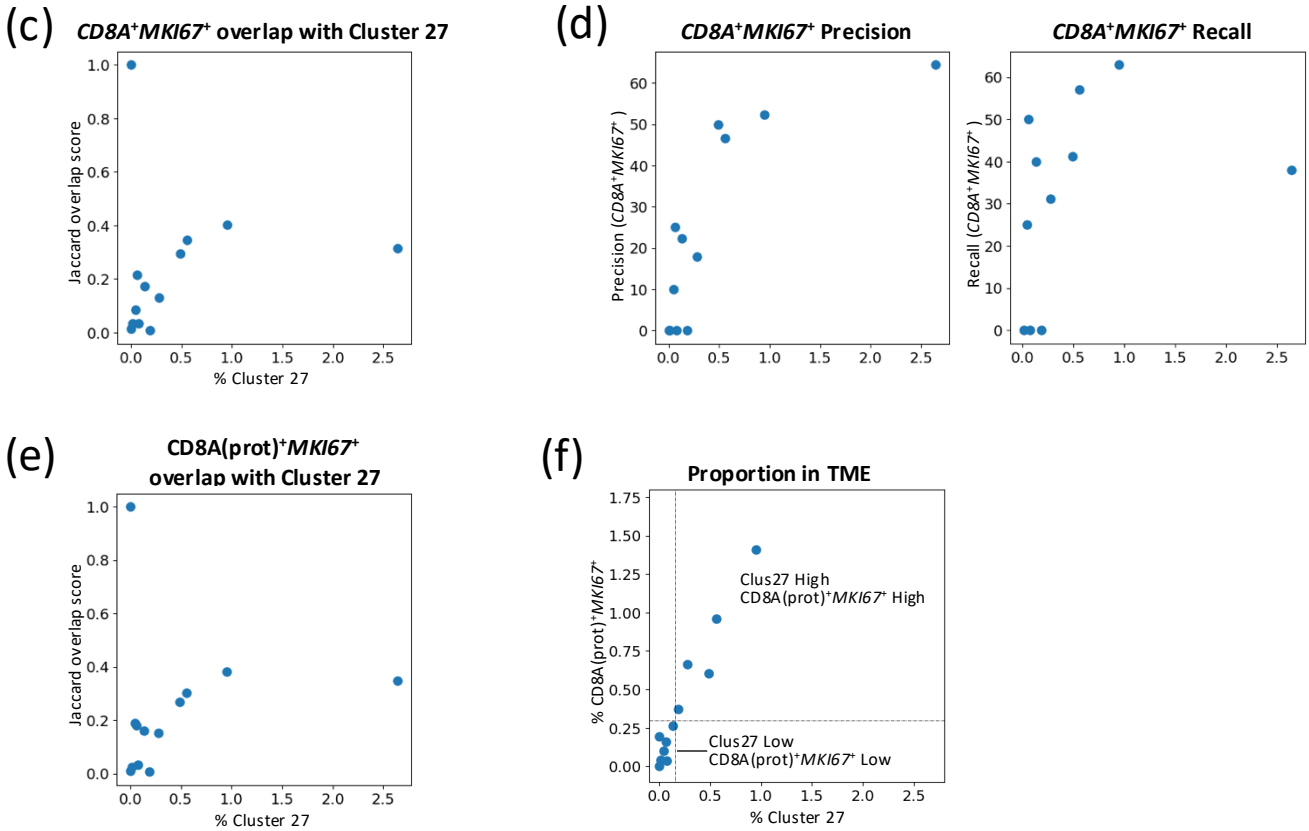
Supplementary Figure 3 (accompanies Figure 3). (a) Clustering of discovery cohort at even finer resolution (54 clusters). The sole *CD8A*⁺*MKI67*⁺ cluster (red box) consists of cells from the original Cluster 27, demonstrating that Cluster 27 cells comprise the only subpopulation of *CD8A*⁺*MKI67*⁺ cells.

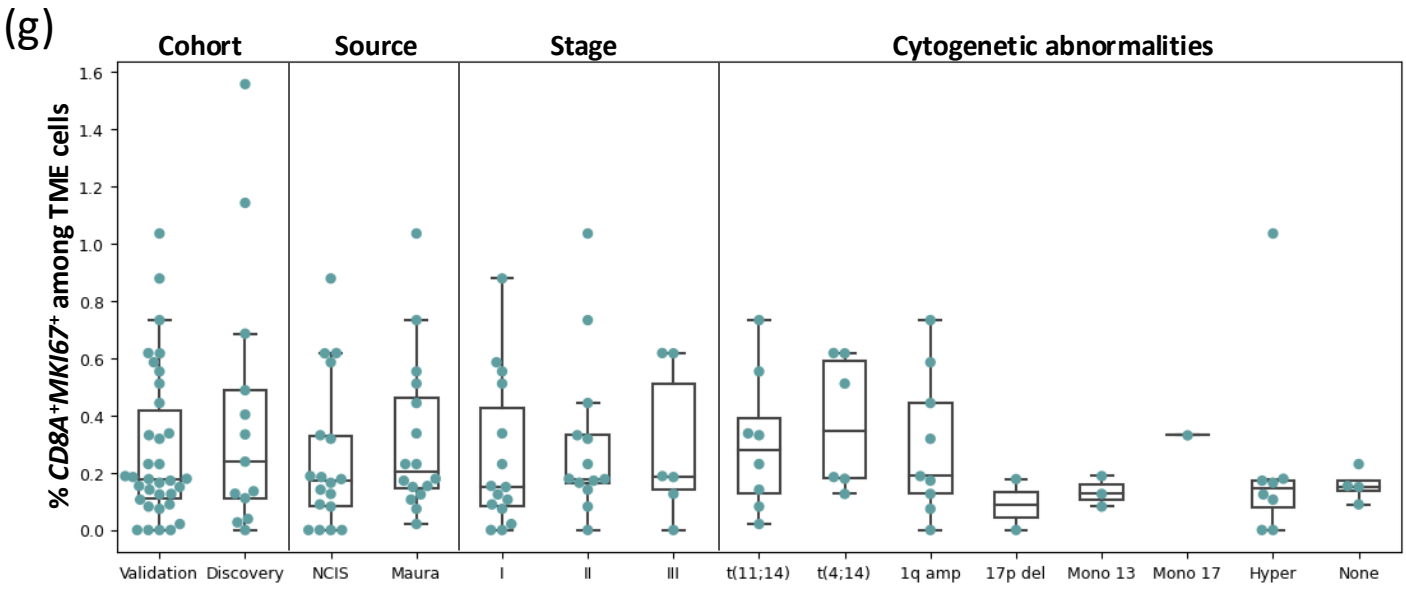


Supplementary Figure 3 (continued). (b) UMAPs showing Cluster 27 cells (left), *CD8A*⁺*MKI67*⁺ cells (middle), and CD8A(prot)⁺*MKI67*⁺ cells (right), among representative Clus27 High and Clus27 Low samples. The signature-positive cells (*CD8A*⁺*MKI67*⁺ and CD8A(prot)⁺*MKI67*⁺ cells) mostly overlap with Cluster 27 cells, albeit with some false positives scattered among other cell types. In Clus27 Low samples, a larger proportion of signature-positive cells constitute false positives, contributing an outsized detrimental effect on the overlap.

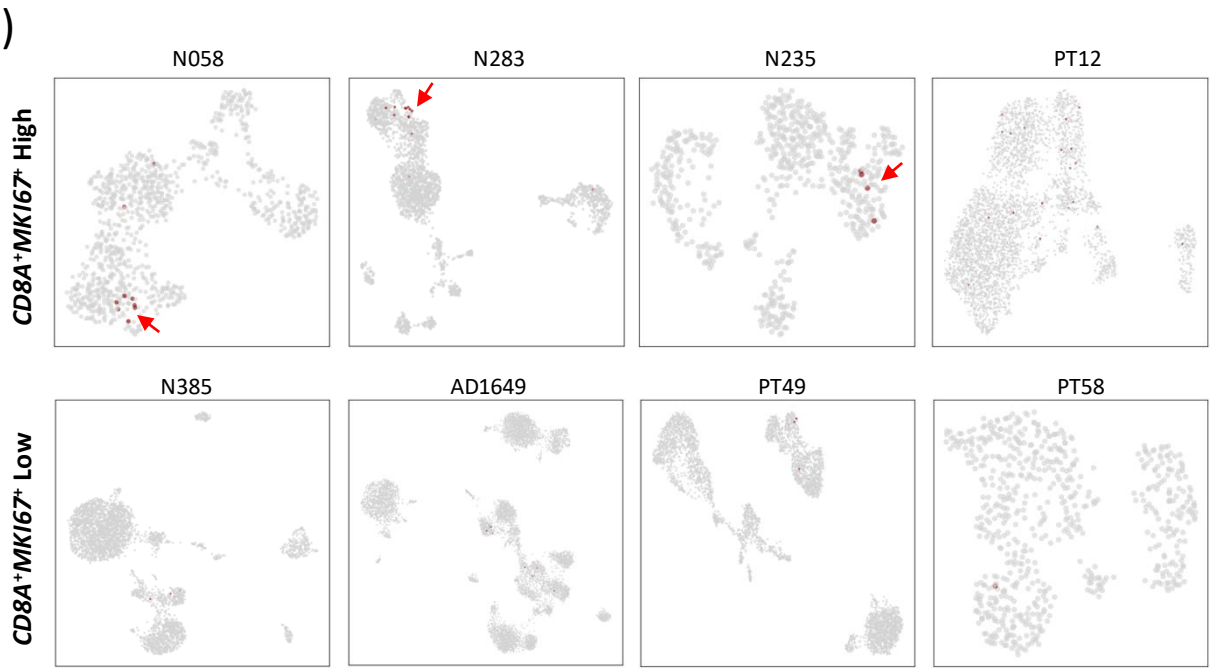


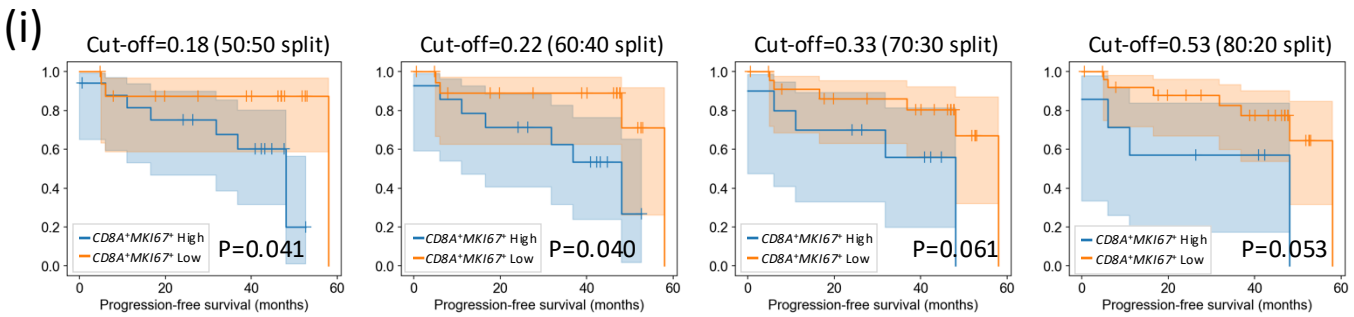
Supplementary Figure 3 (continued). (c) Jaccard overlap score between *CD8A*⁺*MKI67*⁺ cells and Cluster 27 cells in each sample, as a function of the percentage of Cluster 27 cells in the TME. Samples with lower percentage of Cluster 27 cells have lower overlap score. **(d)** Precision and recall of *CD8A*⁺*MKI67*⁺ signature in detecting Cluster 27 cells in each sample, as a function of the percentage of Cluster 27 cells in the TME. Samples with lower percentage of Cluster 27 cells have lower precision and recall. **(e)** Same as (c) but with protein signature CD8A(prot)⁺*MKI67*⁺. **(f)** Percentage of CD8A(prot)⁺*MKI67*⁺ cells in TME correlates with percentage of Cluster 27 cells in TME, with perfect overlap between patient groupings by Clus27 and patient groupings by CD8A(prot)⁺*MKI67*⁺ status. **(g)** Percentages of *CD8A*⁺*MKI67*⁺ cells among TME cells in each sample of the validation cohort, stratified by data source, R-ISS stage, and cytogenetic abnormalities.

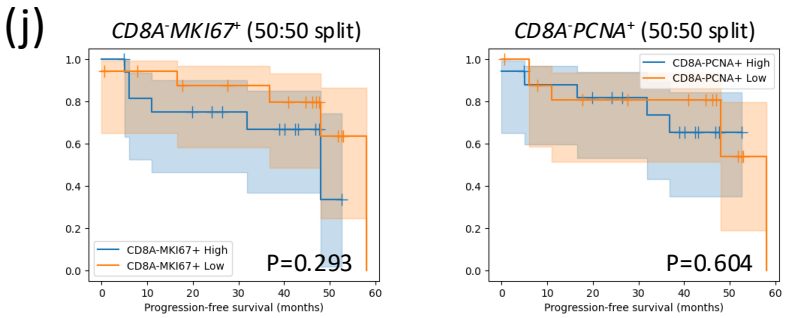




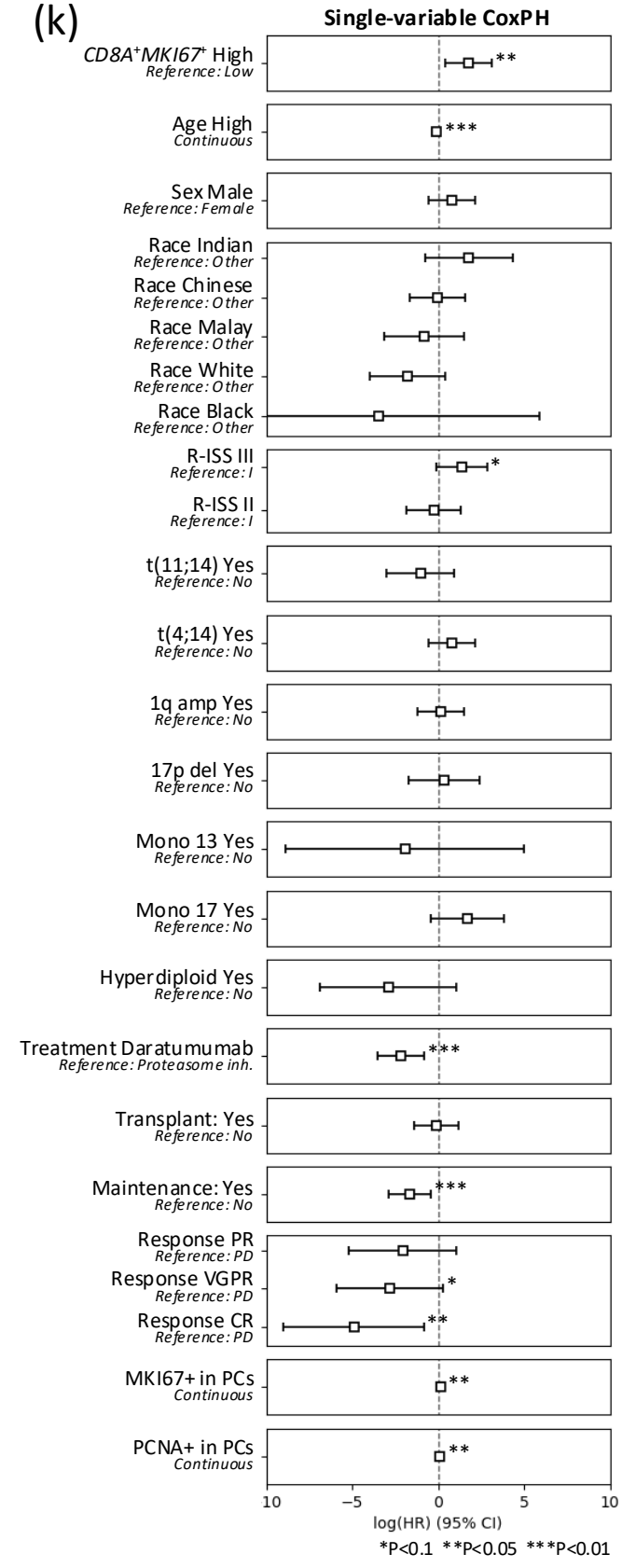
Supplementary Figure 3 (continued). (h) UMAPs showing *CD8A*⁺*MKI67*⁺ cells in representative *CD8A*⁺*MKI67*⁺ High and Low samples. The *CD8A*⁺*MKI67*⁺ High samples have more *CD8A*⁺*MKI67*⁺ cells that are clustered together, whereas the *CD8A*⁺*MKI67*⁺ Low samples have more *CD8A*⁺*MKI67*⁺ cells that are scattered randomly. **(i)** KM survival analyses of *CD8A*⁺*MKI67*⁺ High and Low groups with different cut-offs to split patients into groups. **(j)** Negative control KM survival analyses of *CD8A*⁻*MKI67*⁺ and *CD8A*⁻*PCNA*⁺ High and Low groups. A range of cut-offs were tested and none showed association with PFS; only representative analyses showed here.

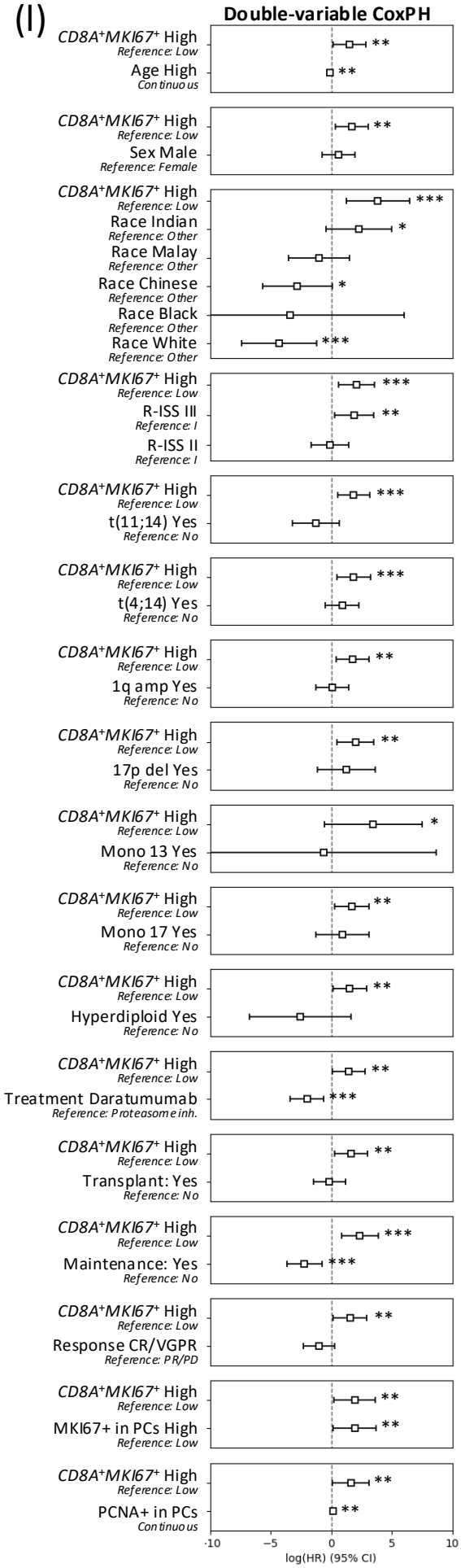




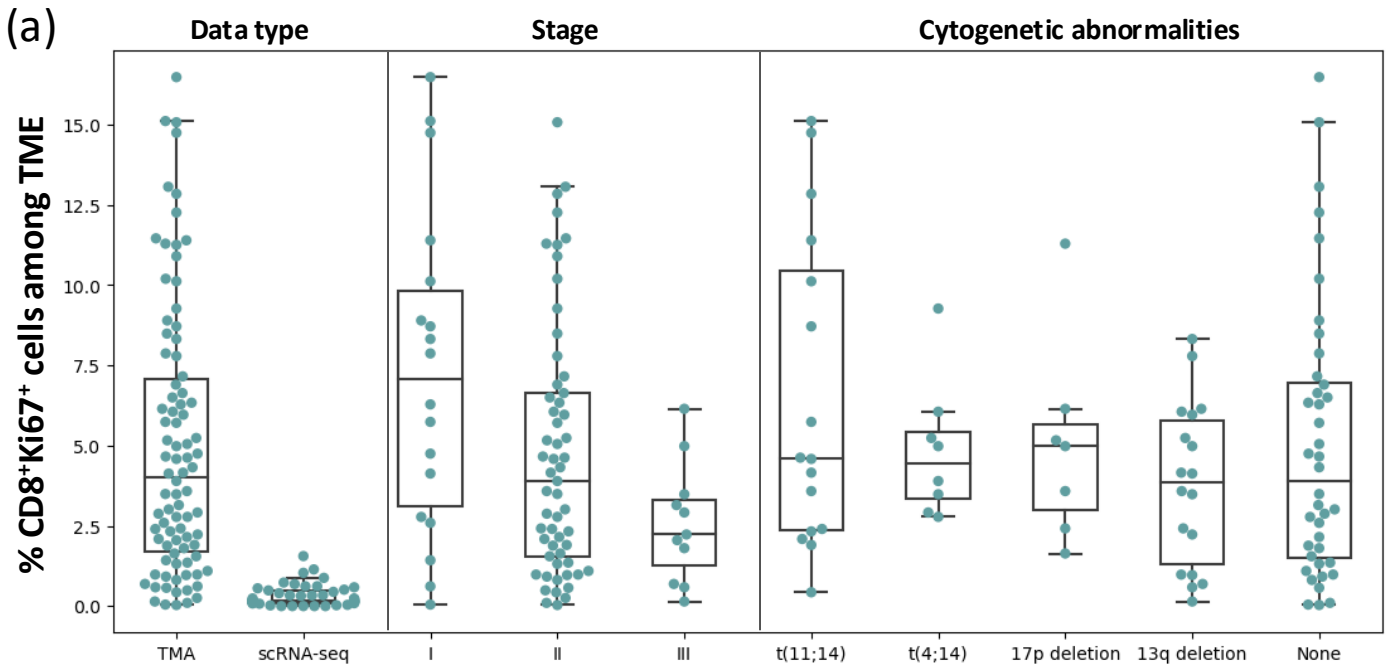


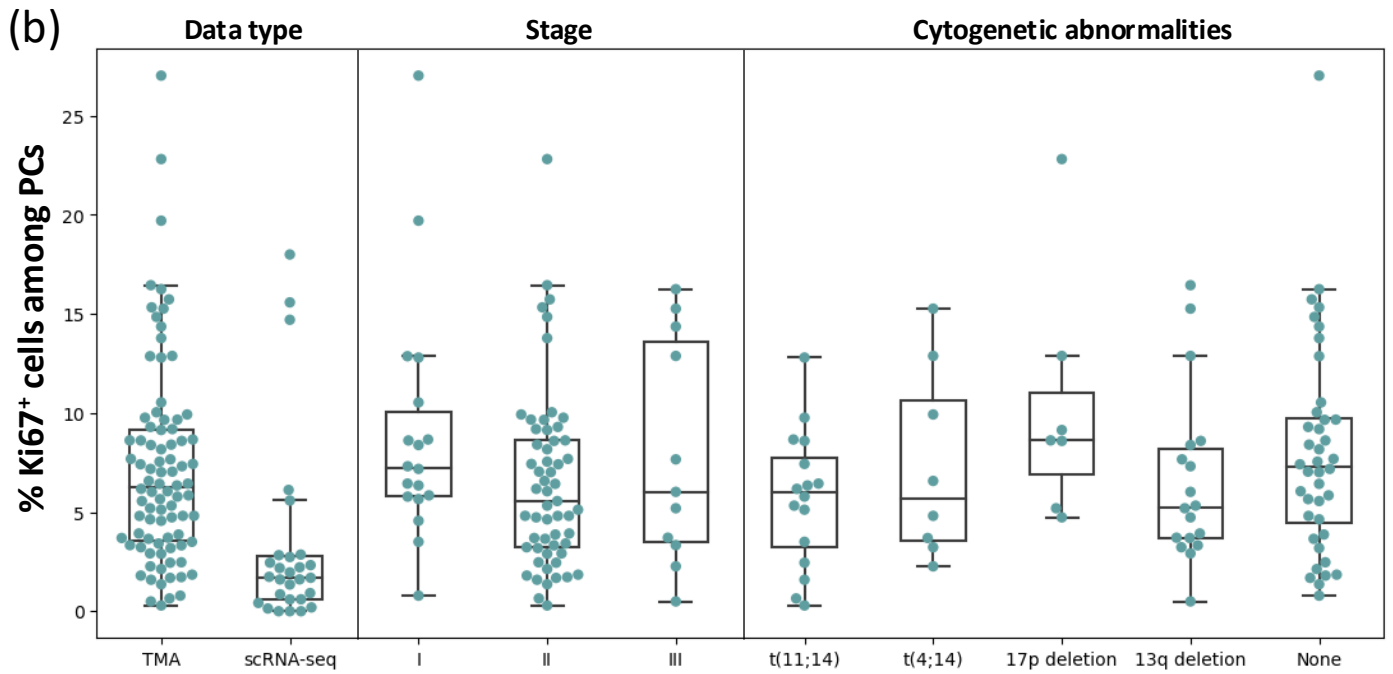
Supplementary Figure 3 (continued). (k) Single-variable and **(l)** double-variable survival analyses with the CoxPH model, for available covariates. Each covariate was evaluated with *CD8A⁺MKI67⁺* status for the double-variable analyses. Covariates with P<0.1 in single or double-variable analysis were selected for multivariable analysis.

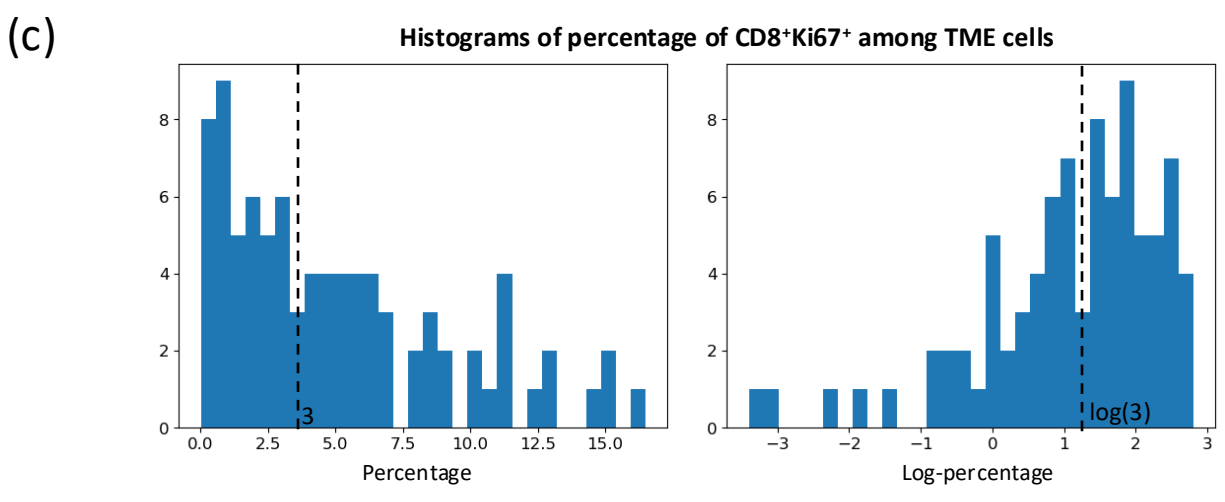




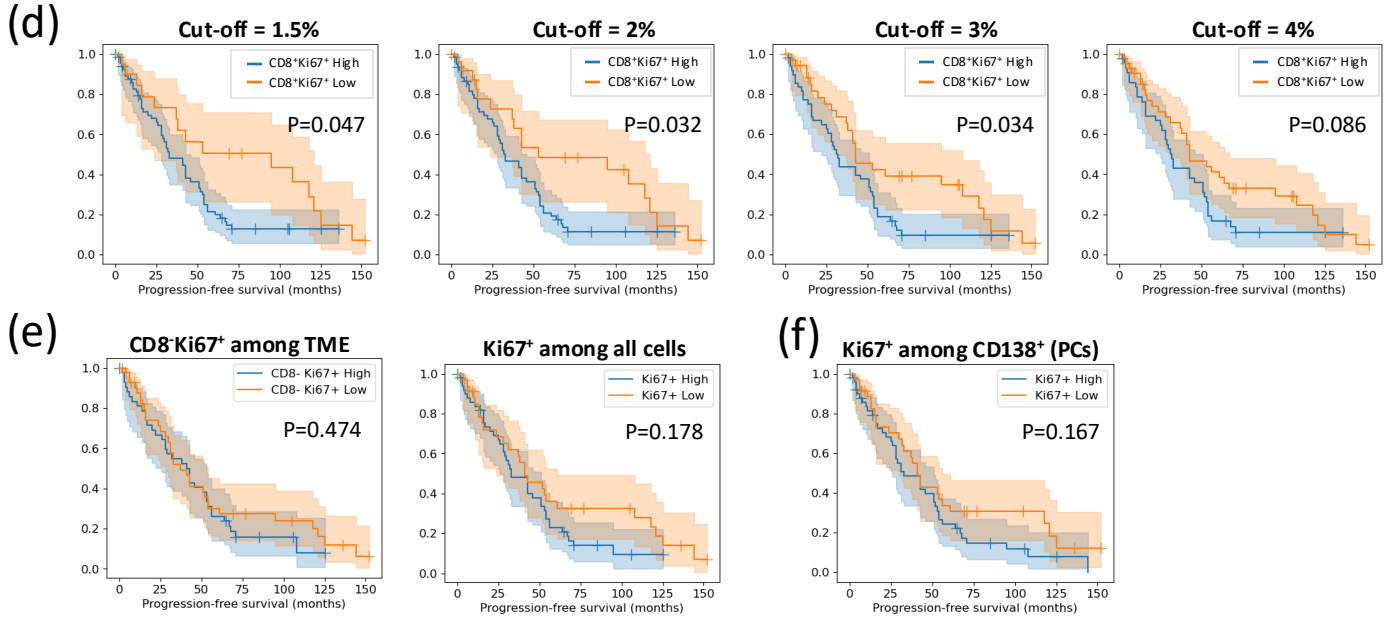
Supplementary Figure 4 (accompanies Figure 4). (a) Percentages of CD8⁺Ki67⁺ cells among TME cells, and **(b)** Ki67⁺ cells among PCs, in each sample of the TMA cohort stratified by R-ISS stage and cytogenetic abnormalities. **(c)** Histogram of CD8⁺ Ki67⁺ cells among TME cells, in percentage and log-percentage scales.

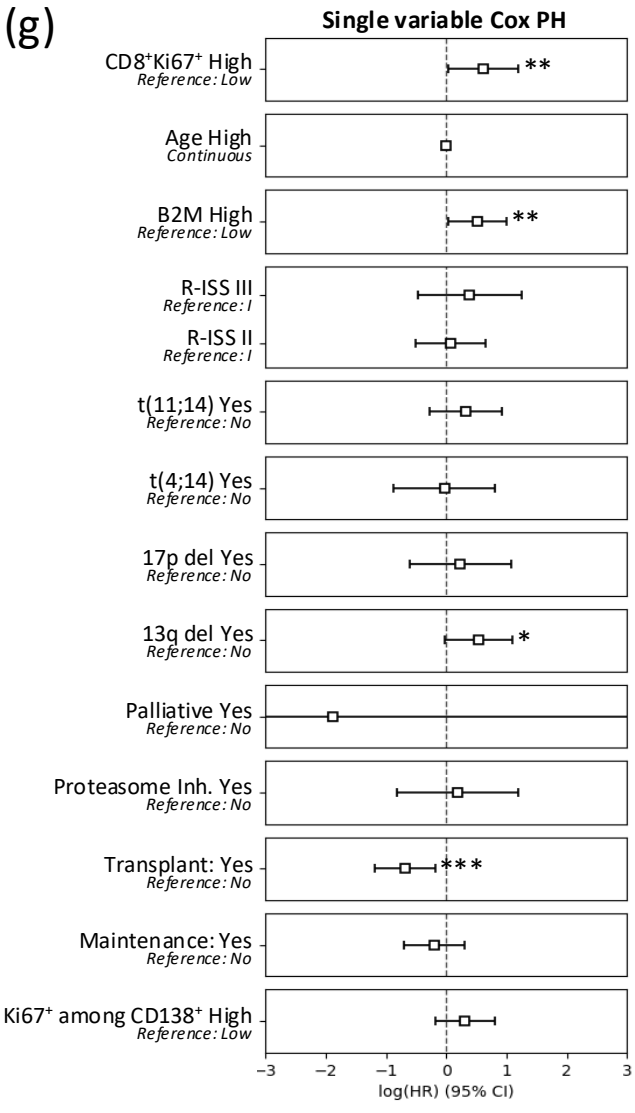






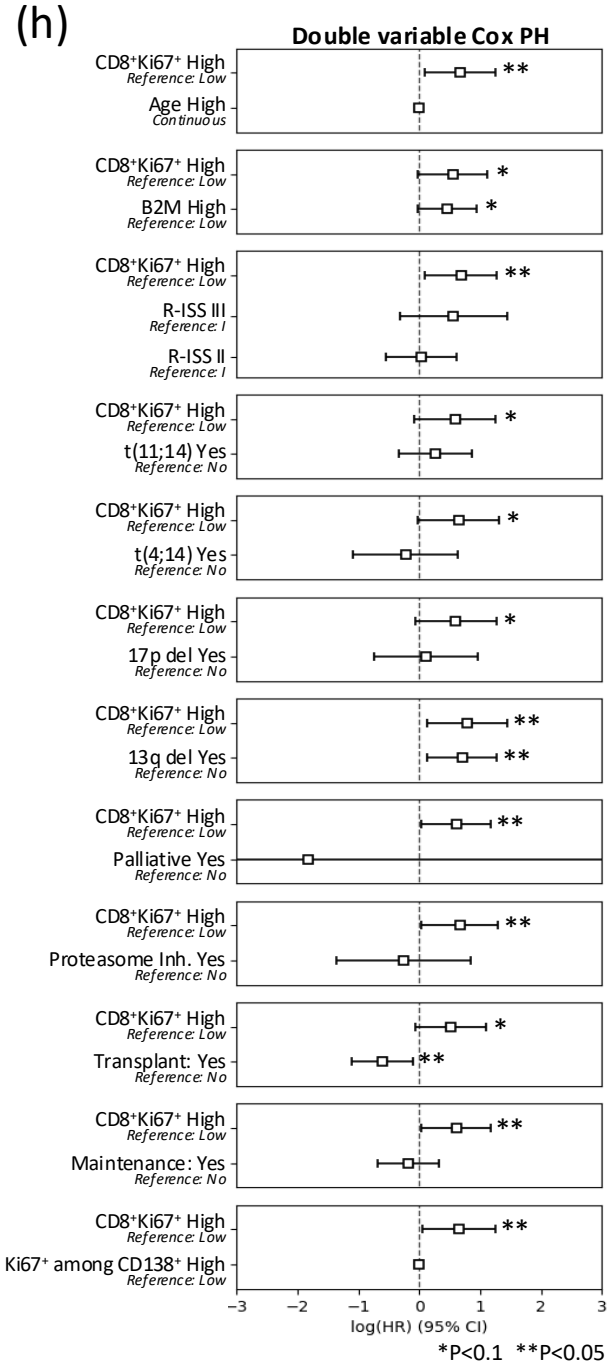
Supplementary Figure 4 (continued). (d) KM survival analyses of CD8⁺Ki67⁺ High and Low groups with different cut-offs to split patients. **(e)** Negative control KM survival analyses of CD8⁻Ki67⁺ among TME cells, and Ki⁺ among all cells, show that neither are associated with PFS. A range of cut-offs were tested and none showed association with PFS; only representative analyses showed here. **(f)** KM survival analyses of the PC proliferative index (Ki67⁺ among PCs).

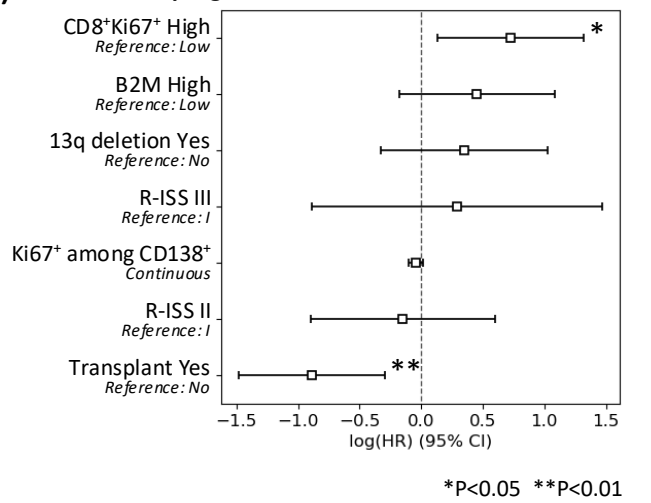


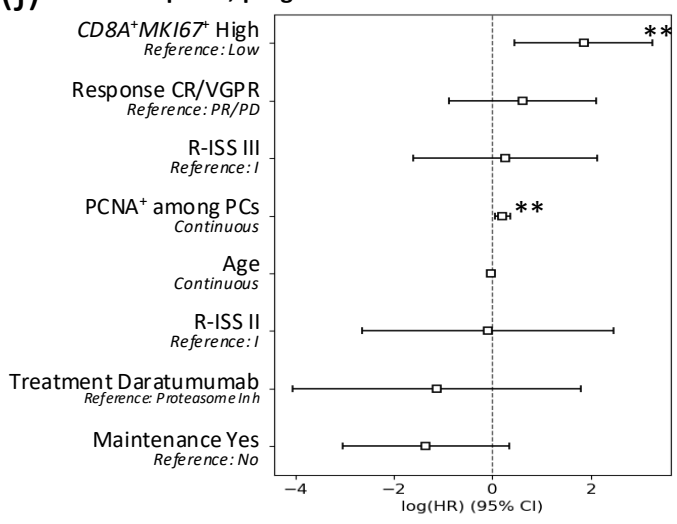


*P<0.1 **P<0.05 ***P<0.01

Supplementary Figure 4 (continued). (g) Single-variable and **(h)** double-variable survival analyses with the CoxPH model, for available covariates. Each covariate was evaluated with $CD8^+Ki67^+$ status for the double-variable analyses. Covariates with P<0.1 in single or double-variable analysis were selected for multivariable analysis. **(i)** CoxPH forest plots for multivariable survival analyses of prognostic markers ($CD8^+Ki67^+$ status, R-ISS, PC proliferative index) and other relevant covariates in the TMA cohort, and **(j)** Combined scRNA-seq data (pooled discovery and validation cohorts). Relevant covariates were selected as those exhibiting association with PFS (CoxPH P<0.1) either individually or together with $CD8^+Ki67^+$ status.

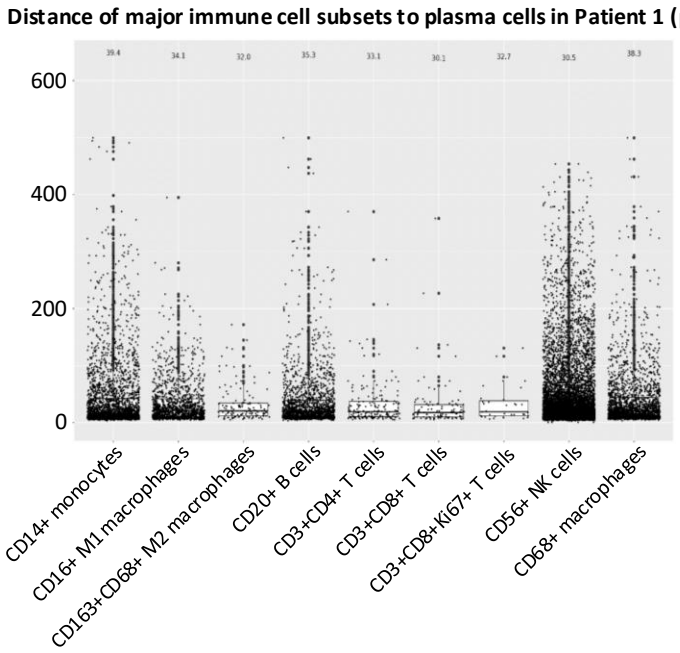




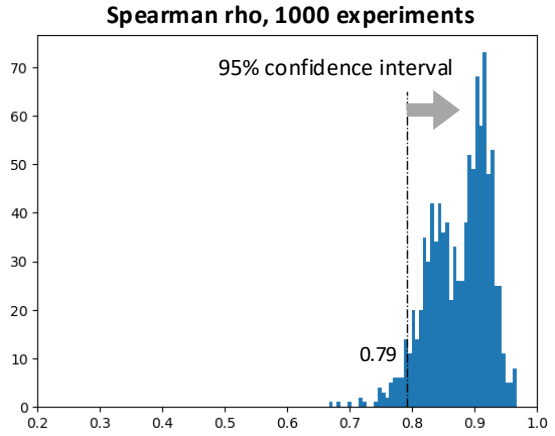


Supplementary Figure 5 (accompanies Figure 5). (a) Distances between proliferative CD8⁺ cells, major immune cells types and myeloma PCs within 500µm, for Patient 1. **(b)** Distribution of Spearman rho correlation coefficients between the ordering of patients by percentage of *CD8A⁺MKI67⁺* among the TME cells from the original data, and from resampled TME cells in 1000 bootstrapping experiments. The ordering of patients in the resampled data were highly correlated with the original data, with rho>0.79 95% of the time. **(c)** Logrank P values for survival analyses of *CD8A⁺MKI67⁺* status from the resampled data. *CD8A⁺MKI67⁺* status was largely associated with worse PFS in the resampled data, with P<0.075 95% of the time. **(d)** Expression profiles of marker genes corresponding to that of Figure 1D of Li et al., 2019. Red boxes indicate genes whose expression appears different between our proliferative CD8⁺ cells (green boxes) and the dysfunctional CD8 T cells of Li et al. 2019.

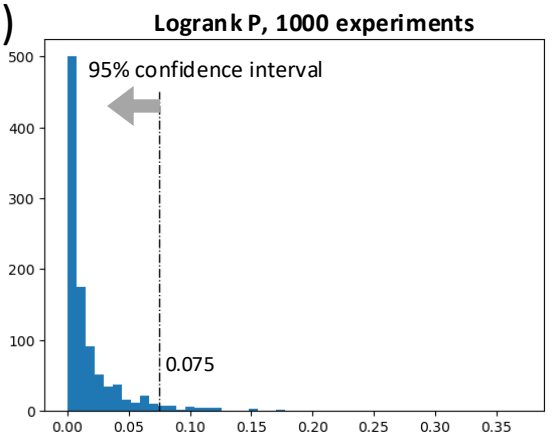
(a)



(b)



(c)



(d)